



Guide à l'usage des dirigeants qui souhaitent doper la croissance de leur entreprise avec une IA moderne

Rentabilisez vos investissements en IA
de manière spectaculaire à l'ère des plateformes



Résumé

Dans tous les secteurs d'activité, les disrupteurs ont introduit de nouvelles références en termes d'expérience client, de délais de mise sur le marché et d'innovation. L'utilisation de l'intelligence artificielle (IA) a atteint un point d'inflexion où les organisations les plus en pointe obtiennent des résultats spectaculaires, font évoluer le marché et se démarquent de leurs concurrents. Au cœur de l'IA se trouvent plusieurs leviers stratégiques : l'automatisation, la prédiction et l'optimisation. La capacité de votre organisation à automatiser les tâches répétitives, à prédire ses résultats et à optimiser ses ressources est essentielle à votre développement. De fait, les entreprises à forte croissance répondent aux demandes du marché – améliorer l'expérience client, fournir des produits et des services plus rapidement, rationaliser les opérations et capitaliser sur l'écosystème – et aux exigences en termes de conformité et de gestion des risques.

Cet article analyse les aspects suivants :

- Caractéristiques des investissements en IA des leaders à forte croissance
- Pourquoi une plateforme de données et d'IA est nécessaire
- Ce qu'une plateforme de données et d'IA doit proposer : automatisation, prédiction et optimisation
- Avantages de créer et de dimensionner l'IA avec confiance et transparence

En lisant cet article, vous verrez comment les leaders de l'industrie tirent parti de l'IA et vous comprendrez l'importance des plateformes ainsi que les avantages d'une plateforme de données et d'IA. De plus, ce guide présente les mesures que vous pouvez prendre et examine les stratégies qui aident votre entreprise à réussir.

« La gouvernance des données et notre gestion des modèles d'IA – la façon dont ils sont validés et utilisés – sont des sujets qui relèvent aujourd'hui du Conseil d'administration. Tout comme l'utilisation éthique des données. »¹

Directeur des opérations dans le secteur bancaire
Pays-Bas

Les opportunités et les défis de l'IA

Les entreprises utilisent l'IA pour prédire leurs résultats, rationaliser leurs opérations, gagner en efficacité, se protéger contre les cybermenaces et la fraude, et identifier de nouvelles opportunités de marché. Forts de ces prédictions, les chefs d'entreprise peuvent garder une longueur d'avance sur la concurrence et anticiper les fluctuations du marché. De plus, les dirigeants doivent tenir les objectifs fixés par les actionnaires, tout en modifiant radicalement les processus, les technologies et les organisations lors de la mise en place de l'IA.

Enfin, les conseils d'administration sont particulièrement attentifs aux données et à la gouvernance concernant les modèles d'IA. [L'IBM® Institute of Business Value a publié une étude¹](#) intitulée « Build Your Trust Advantage – Leadership in the era of data and AI everywhere », qui montre une dégradation rapide de la confiance des clients dans les marques et les institutions. Les clients exigent une transparence des données concernant les produits et les services, et veulent l'assurance que leurs données personnelles seront sécurisées et utilisées correctement.



Les entreprises à forte croissance sont gagnantes avec l'IA

Selon le cabinet Forrester Research, les entreprises à forte croissance investissent massivement dans l'IA. Plus de 50 % des personnes interrogées par Forrester tablent sur un retour supérieur à 5 fois leur investissement dans l'IA². Pour remettre cette information en perspective, sachez que les leaders qui ont investi 10 millions de dollars peuvent espérer un retour sur investissement (RSI) de 60 millions de dollars. Ces leaders investissent également deux fois plus dans les données et l'analyse, et 2,5 fois plus dans les plateformes d'IA et d'apprentissage automatique (AA) que les entreprises à faible croissance.

Forrester Research a également constaté que les sociétés qui investissent dans les « data scientists » ayant des compétences spécifiques – création de modèles prédictifs, d'apprentissage automatique, d'apprentissage profond, de traitement du langage naturel, de la vision des ordinateurs, etc. – se développent plus rapidement que les autres.

Pour certaines applications de l'IA, comme les RH, l'évaluation des prospects ou la gestion des dépenses, notamment la détection de la fraude, les organisations préfèrent acheter des solutions d'IA prêtes à l'emploi, selon Forrester.

- 46 % achètent des solutions prêtes à l'emploi, intégrant une technologie d'IA pour certaines applications.
- 20 % développent leur solution d'IA en interne.

Concrètement, vous avez besoin d'une plateforme de données et d'IA qui vous permette d'acheter et/ou de créer en fonction de vos besoins.

Critères à prendre en compte lors de l'évaluation d'une plateforme de données et d'IA

La science des données – une discipline qui aide une entreprise à identifier des éléments pertinents, à prédire les résultats et à simplifier les décisions – est un élément clé dans l'utilisation de l'IA. Grâce aux nouvelles informations, tendances et autres découvertes dans les données, votre entreprise peut anticiper l'avenir et simplifier les décisions. Avec les bonnes offres et les bonnes approches, vous pouvez optimiser les actions, en recherchant les meilleurs résultats possibles selon les scénarios retenus. Un tel potentiel justifie le recours à une plateforme qui vous aide à concrétiser vos idées et à tirer parti de la science des données progressivement.

Mais, qu'est-ce qu'une plateforme ? Une plateforme est « une infrastructure qui favorise les interactions créatrices de valeur entre les participants. Elle fournit un environnement ouvert et définit les conditions de gouvernance de ces interactions³. » Pour réussir à moderniser votre entreprise avec l'IA, choisir la bonne plateforme est un impératif stratégique.

Créer une base pour gérer les talents, les données et la confiance

Votre objectif consiste à transformer les processus de prédiction et d'optimisation en des workflows intelligents et d'innovation interactifs. Pour ce faire et généraliser l'IA dans vos opérations, votre personnel a besoin d'une intégration simple dans une plateforme de données et d'IA, qui contribue à automatiser le développement. De plus, vous devez trouver des solutions pour réemployer vos investissements technologiques existants dans votre nouvelle plateforme, au lieu d'ajouter des outils ponctuellement.

Talents : Les compétences et le personnel sont des éléments essentiels à la réussite de l'IA. Comme le montre l'article « 2019 AI Predictions » de PwC, votre main-d'œuvre doit maîtriser l'IA. Cela passe par des initiatives continues de requalification et de perfectionnement de vos équipes. De plus, à mesure que les descriptions de poste évoluent, vous devez repenser votre structure organisationnelle pour former votre personnel aux nouvelles compétences⁴.

Données : Les données sont le moteur de la performance du modèle. Les modèles d'IA reposent sur les données. La qualité des données de production détermine le niveau de performance d'un modèle. Une plateforme doit donc permettre d'accéder à un flux continu de données. Elle doit également détecter et limiter la perte inévitable de précision, qui survient lorsque vos modèles rencontrent des données de production différentes de celles sur lesquelles ils sont formés. Et elle doit fournir une trace auditable des modèles et des données utilisées tout au long du cycle de vie de l'IA.

Confiance : Votre équipe doit pouvoir montrer comment détecter et limiter les biais dans vos modèles d'IA, et comment expliquer chaque résultat. Votre plateforme doit aussi être capable d'évaluer les résultats en fonction d'indicateurs de performance métier. Et elle doit intégrer la confiance et l'explicabilité pour vous aider à dimensionner et à soutenir vos initiatives en faveur de l'IA.

Intégrer les prédictions et optimisations de l'IA rapidement dans les applications

Une plateforme de données et d'IA doit également assurer l'intégration des informations générées par les modèles d'IA, dans vos applications modernes. La plupart des entreprises ont déjà beaucoup investi dans le développement des applications. Une plateforme de données et d'IA souple et ouverte peut aider votre équipe commerciale et vos développeurs à élaborer des opérations de modèle (ModelOps). La fonction ModelOps peut collaborer en toute transparence avec la fonction DevOps (développement) pour perfectionner vos applications modernes avec l'IA.

Automatiser la gestion du cycle de vie de l'IA

Historiquement, les investissements en science des données et en IA se sont concentrés sur l'analyse prédictive et l'apprentissage automatique pour répondre aux questions métier ou automatiser un jeu restreint de processus. Mais, la plupart des leaders cherchent maintenant à élargir le champ d'application de l'IA. En d'autres termes, votre plateforme doit être conçue pour opérationnaliser et automatiser la gestion des modèles et des outils de bout en bout dans votre entreprise.

L'automatisation aide votre équipe à se recentrer sur les activités à plus forte valeur ajoutée qui démarquent davantage de la concurrence. Privilégiez une plateforme capable d'automatiser les étapes suivantes :

- Préparation des données
- Ingénierie des fonctionnalités
- Sélection des algorithmes d'apprentissage automatique
- Optimisation des hyperparamètres pour choisir le meilleur modèle possible d'apprentissage automatique

Cette suite d'étapes doit être pilotée par un système d'IA qui opte pour la voie la plus prometteuse à chaque étape du processus. Ainsi, vous utilisez l'IA pour créer de l'IA. C'est ce que propose AutoAI, une fonctionnalité mise en œuvre par IBM Research^{TM5}.

Optimiser les décisions en fonction des prédictions

Certains leaders à forte croissance savent gérer les différentes utilisations de l'IA et améliorer la prise de décision grâce à l'IA. Vous pouvez combiner les informations générées par l'IA et l'orchestration de l'IA avec le talent de vos équipes. Pour aider votre entreprise à optimiser son RSI, une plateforme de données et d'IA doit exploiter les résultats des prédictions et s'en servir pour recommander des actions.

Une plateforme moderne de données et d'IA doit faciliter la sélection et la modification des données en fonction du problème d'optimisation de votre équipe. Équipée d'une interface en langage naturel, la plateforme permet aux membres de votre équipe d'exécuter des modèles d'optimisation, mais aussi de créer et partager des rapports avec des diagrammes de Gantt, des plannings, des plans de ressources et les allocations en termes d'offre et de demande. Intégrer l'optimisation des décisions dans la plateforme de données et d'IA simplifie l'application de l'analyse prescriptive aux prédictions. C'est ce que l'on appelle l'intelligence décisionnelle⁶.

Doper la productivité en perfectionnant et en requalifiant la main-d'œuvre

Les équipes spécialistes en science des données dans les sociétés à forte croissance sont très diverses. Personne ne peut être expert dans les différents domaines de l'IA, comme le calcul, la gestion des données, les utilisations métier, la science de la décision, etc. L'IA nécessite la collaboration de plusieurs équipes ayant différentes compétences et différents points de vue.

Votre plateforme doit donc prendre en charge les approches programmatiques et visuelles de la création de modèles, en utilisant des outils visuels de science des données, des ressources Open Source et des langages de programmation tels que Python ou R. De plus, elle doit aider les « data scientists » visuels à combiner rapidement du texte et des données structurées pour extraire des informations cachées et les utiliser à des fins de prédiction et d'optimisation.

Exploiter vos investissements dans le Big Data

Une plateforme doit impérativement rapprocher l'analyse du Big Data pour extraire des informations plus précisément et avec plus de sécurité. Par exemple, une plateforme moderne de données et d'IA peut sécuriser l'accès aux données et déléguer les exécutions à un cluster Hadoop, sans déplacer un volume important de données. Elle peut également améliorer le développement de modèles et la génération d'informations grâce aux investissements déjà réalisés dans Hadoop, tout en exploitant l'environnement de calcul distribué pour le dimensionnement et la haute disponibilité.

Promouvoir la confiance et la transparence dans les modèles d'IA

Instiller la confiance dans les modèles de données et d'IA est essentiel. Pour réussir, vous devez suivre et mesurer les résultats de l'IA tout au long de son cycle de vie. Vous devez également pouvoir vérifier que vos modèles sont justes, explicables et conformes, quel que soit l'endroit où ils ont été développés ou le cloud sur lequel ils s'exécutent. Votre entreprise et les autorités de régulation exigeant de limiter les biais potentiels et d'expliquer les résultats, ces fonctionnalités doivent faire partie d'une plateforme moderne de données et d'IA.

Fournir un écosystème d'outils Open Source incomparables dans n'importe quel cloud

Votre équipe est partout, tout comme les données. Pour tirer parti des innovations à travers le monde, vous devez amener vos modèles là où vos données se trouvent. Votre plateforme de données et d'IA doit être ouverte et prendre en charge les modèles et les données sur plusieurs clouds, tout en tirant parti d'écosystèmes dynamiques. Elle doit aussi vous permettre de maîtriser le coût et le risque liés à la nécessité de déplacer des données, ce qui peut poser d'éventuels problèmes réglementaires et juridiques.

Par ailleurs, la plateforme devrait aussi permettre de lancer vos projets d'IA avec des accélérateurs industriels et proposer des applications préconfigurées avec des conditions commerciales et des artefacts de science des données prédéfinis⁷.

Combiner les possibilités de création et d'achat dans des modèles de licence souples

Du point de vue technologique, choisir entre la création et l'achat n'est pas facile. La plupart des entreprises à forte croissance choisissent quand acheter et quand créer en termes de science des données et d'IA. Vos besoins peuvent évoluer tout au long de votre cycle de vie. Vous devez donc trouver une plateforme de données et d'IA souple, capable de prendre en charge des applications d'IA prêtes à l'emploi et des outils professionnels qui allient des technologies Open Source et propriétaires.

Mettre en place une pratique d'IA agile avec des outils intégrés

Pour tenir les promesses de la science des données et de l'IA, votre entreprise doit mettre en œuvre une IA agile. Suivez les bonnes pratiques qui sont le fruit de centaines d'implémentations réussies d'IA⁸.

Le respect des principes techniques de base, tels que ceux énoncés ici, vous aidera à former d'autres chefs d'entreprises sur l'IA. Après avoir expliqué comment les [sociétés à forte croissance utilisent l'IA](#), vous pouvez insister sur le fait que la collaboration permet d'obtenir des gains de productivité significatifs aux niveaux individuel et organisationnel. Le soutien de la direction à un centre d'excellence (CoE) en IA contribue à assurer un retour élevé sur investissement.

Watson Studio Premium for IBM Cloud Pak for Data

Watson™ Studio Premium for IBM Cloud Pak™ for Data vous aide à valoriser plus rapidement vos investissements en IA. Watson Studio Premium comprend IBM Decision Optimization, SPSS® Modeler et Hadoop Execution Engine.

Cette puissante association aide une entreprise à :

- Simplifier les décisions en fonction des prédictions obtenues par la plateforme de données et d'IA
- Employer des spécialistes des données et de l'analyse partout
- Tirer parti de vos investissements existants dans le Big Data

La plateforme permet aux organisations de prédire et d'optimiser les résultats au sein d'un environnement unifié.

Fonctionnant sur Red Hat® OpenShift® et déployable en quelques heures, IBM Cloud Pak for Data peut facilement être complété avec un éventail croissant de microservices IBM et tiers. IBM Cloud Pak for Data s'exécute sur n'importe quel cloud, permettant aux entreprises d'intégrer plus facilement leurs analyses et leurs applications pour dynamiser l'innovation. Votre entreprise peut créer et déployer des modèles d'IA et d'AA, finaliser plus rapidement des projets de science des données et mettre en place un cycle de vie automatisé de bout en bout sur cette plateforme moderne d'IA. IBM Cloud Pak for Data prend totalement en charge les environnements multicloud, tels qu'Amazon Web Services (AWS), Azure, Google Cloud, IBM Cloud™ et les clouds privés. AutoAI et Watson OpenScale™ sont disponibles dans la version de base d'IBM Cloud Pak for Data.



Utilisations transversales de l'IA

Rendre les centres d'appels plus efficaces avec des agents conversationnels capables de gérer les pics imprévisibles d'appels, d'e-mails, de SMS et de messages instantanés des clients

Augmenter la vente croisée et la vente incitative avec des offres et des recommandations en temps réel personnalisées

Accroître la fidélité des clients en anticipant l'attrition et en proposant des mesures de fidélisation

Optimiser les offres en écoutant l'avis des clients et en anticipant les besoins futurs

Améliorer le marketing avec des campagnes ciblées et personnalisées

Minimiser les coûts des stocks et améliorer la gestion des ressources grâce à des prévisions fiables

Augmenter la productivité en affectant les bons employés aux bons postes au bon moment et en générant des prévisions fiables concernant la main-d'œuvre

Réduire les coûts de maintenance en anticipant les défauts avant qu'ils ne surviennent

Maîtriser le risque par une évaluation fiable de la solvabilité des clients

Détecter la fraude en identifiant les comportements suspects

Déverrouiller les nouveaux modèles économiques en répondant aux demandes non prises en compte et en intégrant la prédiction dans les applications modernes

Avantages de créer et de dimensionner l'IA avec confiance et transparence

Vous dirigez votre entreprise de manière à réussir sa transition vers l'IA. La monétisation de l'IA vous oblige à investir dans la bonne plateforme de données et d'IA, une plateforme qui relève les défis de manière transformationnelle en termes de données, de talents et de confiance. Cette plateforme doit intégrer la prédiction, l'automatisation et l'optimisation dans votre implémentation de l'IA. En créant et en dimensionnant l'IA avec confiance et transparence, vous gardez la maîtrise sur les activités suivantes :

- Automatisation de la gestion du cycle de vie de l'IA
- Amélioration de l'intelligence décisionnelle grâce à la prédiction
- Suivi de la performance du modèle
- Promotion de la science des données collaborative sur une plateforme de données et d'IA multicloud unifiée

Les avantages sont substantiels. Votre organisation peut surclasser ses concurrents en termes de chiffre d'affaires, de rentabilité, d'excellence opérationnelle et d'innovation. Vous gérez les données et les modèles d'une manière qui favorise la confiance et qui mobilise les clients et les partenaires commerciaux différemment. De plus, vous donnez confiance dans les modèles de données et d'IA, et vous les utilisez pour améliorer les expériences des clients et des partenaires tout au long de leurs chaînes de valeur pour dynamiser la croissance.

Pour en savoir plus

Découvrez comment IBM Cloud Pak for Data peut transformer votre entreprise.

[Voir comment cette solution fonctionne →](#)

Webinaire : Mémo-guide Gagner avec l'IA

Découvrez comment obtenir un meilleur retour sur vos investissements en IA et ce qui caractérise les leaders dans ce webinaire en trois parties.

[S'inscrire au webinaire →](#)

Le RSI de l'IA d'entreprise

Seth Dobrin, Vice-président d'IBM et directeur des données, explique comment générer un meilleur RSI en intégrant totalement l'IA dans l'entreprise au lieu de la limiter à l'exploration et à la recherche.

[Regarder la vidéo \(2:08\) →](#)

Évaluation de la valeur métier

Les solutions d'IA sécurisées qui améliorent les processus peuvent générer une valeur métier importante. Cette évaluation de la valeur métier de l'IA vous montre tout l'intérêt de cette technologie.

[Effectuer l'évaluation →](#)

Découverte du produit

Pour valoriser l'IA plus rapidement et obtenir de manière automatique des informations et des décisions fiables, vous devez vous assurer que vos données sont de qualité. IBM Cloud Pak for Data peut vous aider. Découvrez le produit pour en savoir plus.

[Découvrir le produit →](#)

Annexe : Utilisations de l'IA par secteur d'activité

Comment les leaders à forte croissance qui investissent dans la création de leurs propres modèles d'IA font-ils pour résoudre leurs problèmes métier ? Examinons quelques exemples par secteur d'activité :

Banque

- Évaluer les risques liés au marché et aux contreparties pour les échanges
- Évaluer le risque du crédit pour les applications de prêt
- Détecter les transactions frauduleuses en temps réel
- Accélérer l'obtention de prêts et le service client tout en maîtrisant le risque lié aux évaluations rapides du risque de crédit

Assurance

- Détecter les réclamations frauduleuses
- Optimiser les devis et les primes en évaluant les risques pertinents pour chaque demandeur
- Prédire les événements météorologiques dangereux pour réduire le nombre de demandes d'indemnisation automobile

Énergie et services publics

- Gérer les réseaux importants d'actifs physiques
- Anticiper l'évolution de la production et de la demande
- Prévoir les coupures de courant avant qu'elles ne surviennent
- Planifier l'offre et la demande
- Rationaliser le processus d'optimisation de l'offre et de la demande

Administration publique

- Détecter les fraudes aux prestations sociales
- Prédire l'évolution de l'utilisation des services publics
- Optimiser la gestion des déchets et les flux de trafic

Industrie manufacturière

- Assurer le bon fonctionnement des lignes de production en modélisant la qualité des produits et en détectant les défauts
- Optimiser la gestion des entrepôts et la logistique
- Développer des capteurs pour véhicules autonomes à l'aide de modèles de ML (Machine Learning)
- Optimiser les opérations de la chaîne logistique

Vente au détail

- Gérer les programmes de fidélité de la clientèle
- Développer la vente croisée et la vente incitative en proposant des recommandations ciblées selon les profils des clients et des modèles de propension sophistiqués
- Prévoir la demande avec une fiabilité accrue

Alimentation

- Automatiser la collecte et l'analyse des données sur la santé des aliments
- Prédire et signaler les crises potentielles pour favoriser une intervention rapide
- Protéger les données sensibles pour permettre aux concurrents de collaborer en toute sécurité

Santé

- Surveiller les flux de données des ECG et des autres dispositifs médicaux
- Prédire l'évolution de l'état de santé d'un patient
- Effectuer des recherches médicales
- Analyser les flux de données des patients en temps réel

Médias et divertissement

- Fournir plus rapidement des analyses plus approfondies sur les audiences de la télévision
- Accélérer l'extraction de données d'audience dans des ensembles plus complexes et plus complets
- Assurer un dimensionnement simple et rapide en fonction de la demande, grâce à des ressources souples
- Se concentrer sur l'activité plutôt que sur la technologie ou les opérations

Enseignement

- Prévoir la réussite scolaire et l'assiduité
- Identifier les élèves qui ont besoin d'un soutien particulier pour atteindre leurs objectifs
- Renforcer les relations avec les mécènes
- Suivre les déplacements des élèves pour réduire l'absentéisme

Services informatiques

- Fournir des informations instantanées sur les pratiques professionnelles d'une société
- Atteindre une efficacité sans précédent grâce à l'automatisation robotisée des processus (RPA)
- Aider les employés à se concentrer sur les activités stratégiques
- Renforcer la satisfaction des clients en améliorant la qualité du service grâce aux modèles de prédiction et d'attrition du comportement des clients

Glossaire

Les **algorithmes** sont des ensembles de règles qui définissent une suite d'opérations applicables à des données pour résoudre un problème particulier.

L'**intelligence artificielle (IA)** désigne la capacité des systèmes informatiques à interpréter les données et à en tirer des conclusions. Ce terme est plus communément utilisé pour décrire les systèmes intégrant des modèles d'apprentissage automatique ou profond. Les techniques d'IA permettent aux ordinateurs de résoudre un large éventail de problèmes auparavant considérés comme insurmontables.

Un **biais** est un problème commun lié à la conception, à la formation et au test de modèles, et pouvant être à l'origine de prévisions erronées. Réduire les biais en analysant et en auditant les modèles pendant l'exécution est une étape de plus en plus importante, à mesure que les entreprises cherchent à généraliser l'IA.

Les **modèles de classification** servent à catégoriser les points de données en les comparant à un ensemble de points de données déjà catégorisés. Le résultat est une valeur discrète, c'est-à-dire une valeur issue d'une liste limitée et non un score. Par exemple, un modèle de classification peut permettre de savoir par « oui » ou « non », si les clients sont susceptibles d'effectuer un achat ou s'ils présentent un risque de solvabilité.

L'**analyse de contenus** consiste à analyser des données non structurées provenant de documents de différents formats, comme du texte, des images, des fichiers audio ou des vidéos. Les techniques d'apprentissage automatique accélèrent considérablement l'analyse de référentiels importants de contenus, qui aurait auparavant nécessité plusieurs centaines ou milliers d'heures de travail.

La **science des données** est une vaste discipline qui réunit certains aspects des statistiques, de l'analyse des données et de l'apprentissage automatique pour résoudre des problèmes métier. Elle permet d'identifier de nouvelles tendances, de générer des prédictions et d'optimiser les décisions.

L'**optimisation décisionnelle** utilise des techniques de pointe en intelligence artificielle et en mathématiques pour résoudre des problèmes de décision mettant en œuvre plusieurs millions de variables, de contraintes métier et de compromis.

L'**apprentissage profond** est une branche de l'apprentissage automatique, qui utilise les réseaux neuronaux avec un nombre important de couches cachées. Ces réseaux très sophistiqués sont employés dans des domaines de pointe de l'apprentissage profond, comme la vision par ordinateur, la traduction automatique et la reconnaissance vocale.

Le **déploiement** est un processus qui consiste à intégrer un modèle dans des applications métier et à exécuter ce modèle sur des données réelles. La création du modèle et son passage par des environnements de test, intermédiaire et de production requièrent une collaboration entre vos « data scientists », vos développeurs d'application et vos équipes informatiques.

La **modélisation prédictive** met en œuvre des techniques statistiques traditionnelles ou des algorithmes d'apprentissage automatique pour créer des modèles et les affiner par des phases de formation et de test sur vos ensembles de données. Ce processus de développement est hautement itératif : il faut former plusieurs dizaines, voire plusieurs centaines, de modèles pour obtenir le niveau de précision souhaité.

L'**explicabilité** contextualise chaque décision, en rendant les modèles d'IA transparents et contrôlables. C'est une caractéristique importante de tout système qui utilise les modèles prédictifs pour fournir des recommandations et faciliter la prise de décision. Un modèle prédictif considéré comme complexe et mystérieux ne gagnera pas la confiance des employés, des autorités de régulation et des clients.

L'**exploration** des données représente une étape importante du processus d'élaboration de modèle. Son but est d'identifier des tendances dignes d'intérêt dans un ensemble de données, de mettre au jour des relations cachées et de proposer des utilisations où la modélisation prédictive pourrait apporter un plus.

L'**analyse géospatiale** consiste à analyser des données géographiques, comme la latitude, la longitude, les codes postaux et les adresses postales. Ce type d'analyse est extrêmement utile pour résoudre de nombreux problèmes pratiques en science des données. Une plateforme moderne de science des données doit faciliter la détection, l'analyse et le calcul d'informations géospatiales, et s'intégrer facilement dans les outils de cartographie pour visualiser les résultats.

En intelligence artificielle, l'**inférence** consiste à appliquer des règles logiques à la base de connaissances pour tirer des conclusions et lever l'incertitude. Ainsi, les utilisateurs obtiennent une prédiction simplifiée, compressée, qui optimise la performance d'exécution.

La **régression linéaire** est un traitement statistique qui utilise une variable indépendante pour expliquer ou prédire une valeur ou un score. Par exemple, ce peut être le nombre de références d'un produit vendues pendant une semaine donnée ou le pourcentage de risque qu'un client ferme son compte.

La **régression logistique** est un traitement statistique employé pour prédire des résultats. Il se distingue de la régression linéaire en ce qu'une variable indépendante n'a qu'un nombre limité – et non illimité – de valeurs possibles. Les utilisateurs emploient la régression logistique lorsque la réponse appartient à certaines catégories, comme un ordre numérique (premier, deuxième, troisième, etc.).

L'**apprentissage automatique (AA)** utilise des techniques statistiques pour générer des algorithmes et des modèles prédictifs sophistiqués à partir de larges ensembles de données, sans passer par une programmation explicite. En général, on débute ce processus itératif en divisant un ensemble des données en deux : un pour la formation et un pour le test. On forme les modèles sur l'ensemble de formation et on teste leurs performances sur l'ensemble de test avec plusieurs dizaines ou centaines de variantes pour évaluer la fiabilité de leurs prédictions. Avec ce processus et en réutilisant les itérations les plus performantes dans la génération suivante de variantes, le modèle apprend progressivement et gagne en exactitude.

La **gestion** des modèles est essentielle pour garantir leur précision au fil du temps. Reformuler les modèles régulièrement en prenant en compte les nouvelles données est essentiel, pour que le développement, l'implémentation, le déploiement et la gestion des modèles soient continus.

La fonction **ModelOps** (opérations de modèle) intègre et déploie un modèle d'IA au développement d'application pour enrichir les cycles de vie de la fonction DevOps et du développement d'application. Le modèle à opérationnaliser peut être un modèle d'apprentissage automatique, d'optimisation décisionnelle ou de transformation des données. Ce n'est pas forcément un modèle d'apprentissage automatique traditionnel, il peut inclure différents modèles.

Le **traitement du langage naturel (NLP)** est une discipline de l'IA qui vise principalement à rendre les ordinateurs capables d'analyser les données textuelles non structurées. On l'utilise essentiellement pour la reconnaissance vocale, la compréhension du langage naturel et l'analyse des sentiments.

Les **réseaux neuronaux** fournissent aux modèles un cadre de formation qui permet des interactions complexes entre de nombreux algorithmes d'apprentissage automatique pour identifier les meilleurs modèles. La structure des réseaux neuronaux artificiels s'inspire de la structure d'interconnexion des neurones dans le cerveau humain et celui d'autres animaux. Des couches relient les neurones artificiels. Les données traversent la structure depuis la couche interne jusqu'à la couche externe en passant par une ou plusieurs couches cachées. Pendant ce parcours, des fonctions mathématiques transforment les données en une prédiction dont vous pouvez évaluer la précision.

Les **logiciels Open Source** sont devenus un paradigme prédominant dans de nombreux aspects de la modélisation statistique et de l'apprentissage automatique. Les langages tels que R, Python et Scala, les architectures comme Apache Hadoop et Spark, ainsi que les infrastructures d'apprentissage automatique telles que TensorFlow et Spark MLlib, sont des acteurs majeurs dans le monde de l'analyse prédictive et de la science des données.

L'**analyse prédictive** utilise les données historiques pour modéliser un domaine ou un problème spécifique, et pour isoler les facteurs clés à l'origine de résultats spécifiques dans le passé. Les modèles créés selon ce processus fournissent des résultats probables à partir de nouvelles données. L'analyse prédictive englobe un large panel de techniques, depuis la modélisation statistique classique jusqu'aux algorithmes d'apprentissage automatique.

Les **modèles prédictifs** sont des algorithmes qui relient une entrée (une donnée, comme un enregistrement de base de données, un texte ou une image) à une sortie (ou prédiction). En général, les sorties sont soit des variables continues (comme un nombre ou un pourcentage), soit des catégories discrètes (comme « oui » ou « non »). Il existe deux types de modèles prédictifs : les modèles de régression et les modèles de classification.

La **préparation des données** est l'une des premières étapes du processus de science des données. La plupart des projets commencent par affiner des ensembles de données afin de vérifier que la qualité est suffisante pour supporter le poids d'une analyse détaillée. Souvent, vos données sources nécessiteront un nettoyage et une transformation dans un format plus adapté à la modélisation et à l'analyse. Si vous créez un modèle d'apprentissage automatique, vous devrez peut-être aussi investir dans l'étiquetage manuel des données en vue d'un apprentissage supervisé.

Les **modèles de régression** sont utiles lorsque votre ensemble de données a plusieurs variables et que vous souhaitez analyser les relations entre celles-ci. Plus précisément, les modèles de régression peuvent montrer comment une variable est susceptible d'évoluer en fonction d'autres variables. La régression linéaire permet de prédire une valeur ou un score. Par exemple, ce peut être le nombre de références d'un produit vendues pendant une semaine donnée ou le pourcentage de risque qu'un client ferme son compte.

La **modélisation statistique** est un domaine des mathématiques qui implique la création de modèles à partir d'hypothèses probabilistes sur un ensemble de données. Les entreprises ont utilisé les modèles statistiques pour analyser les caractéristiques importantes de leurs ensembles de données et identifier des corrélations pouvant servir à classer les données ou à obtenir des prédictions.

L'**apprentissage supervisé** est une méthode permettant de former un modèle d'apprentissage automatique, à l'aide d'un ensemble de données dont le contenu a été correctement étiqueté. Le modèle produit une variable de sortie – en général, une catégorie ou une valeur – dont la précision peut facilement être évaluée en la comparant à l'entrée étiquetée. La régression linéaire, les forêts aléatoires et les machines à vecteurs de support sont des exemples connus d'algorithmes d'apprentissage supervisé, et la plupart des modèles prédictifs sont élaborés avec ces techniques.

Outre la formation, le **test** des modèles prédictifs est essentiel pour déterminer la précision des données dans les processus d'IA. Les modèles prédictifs doivent être testés en continu pour gagner en précision. Si un modèle ne donne pas satisfaction, les analystes doivent identifier la cause première, puis refaire les étapes de formation et de test afin de l'améliorer.

L'**analyse de texte** évalue les contenus non structurés en utilisant les règles linguistiques, le traitement du langage naturel et l'apprentissage automatique. Ce processus analyse les données selon des approches similaires au cerveau humain, mais beaucoup plus rapidement. L'analyse de texte permet d'extraire davantage d'informations et de découvertes dans les contenus non structurés qui représentent environ 90 % de toutes les données.

La **formation** des modèles prédictifs est une phase clé de l'apprentissage automatique, de l'apprentissage profond et d'autres processus d'IA, pour identifier les données utiles. Un modèle formé pour donner des prédictions exactes peut être utilisé pour noter les données en temps réel. La formation des modèles doit être répétée régulièrement pour tenir compte de l'évolution des comportements.

L'**apprentissage non supervisé** est une méthode permettant de former les modèles d'apprentissage automatique à l'aide de données non étiquetées. Souvent, l'objectif est de modéliser et de mettre en valeur des tendances ou des structures intéressantes dans les données. On utilise souvent l'apprentissage non supervisé dans les problèmes de cluster et d'association, notamment pour trouver de nouvelles solutions permettant de segmenter les clients ou d'identifier leurs similarités.

La **visualisation** désigne le processus de représentation graphique des données, souvent sous la forme de tableaux et de diagrammes. Pour comprendre les données, il faut les visualiser. Ce processus est important à la fois pour présenter vos résultats à vos interlocuteurs et pour explorer un nouvel ensemble de données lors des premières phases d'un projet.

© Copyright IBM Corporation 2019

Compagnie IBM France
17 avenue de l'Europe
92275 Bois-Colombes Cedex

Produit aux États-Unis.
Décembre 2019

IBM, le logo IBM, **ibm.com**, IBM Cloud, IBM Cloud Pak, IBM Research, SPSS, Watson et Watson OpenScale sont des marques d'International Business Machines Corp., déposées dans de nombreux pays du monde. Les autres noms de produits et de services peuvent être des marques d'IBM ou d'autres sociétés. Une liste actualisée des marques déposées IBM est accessible sur le web sous la mention « Copyright and trademark information » à l'adresse www.ibm.com/legal/copytrade.shtml.

Red Hat® et OpenShift® sont des marques ou des marques déposées de Red Hat, Inc. ou de ses filiales aux États-Unis et dans d'autres pays.

Ce document est considéré comme à jour à sa date initiale de publication et peut être modifié par IBM à tout moment. Toutes les offres ne sont pas disponibles dans chaque pays où opère IBM.

Il appartient à l'utilisateur d'évaluer et de vérifier le fonctionnement des autres produits ou programmes avec les produits et programmes d'IBM. LES INFORMATIONS CONTENUES DANS CE DOCUMENT SONT FOURNIES « EN L'ÉTAT », SANS AUCUNE GARANTIE EXPRESSE OU TACITE, NOTAMMENT SANS AUCUNE GARANTIE DE QUALITÉ MARCHANDE OU D'ADAPTATION À UN EMPLOI SPÉCIFIQUE, ET SANS AUCUNE GARANTIE OU CONDITION DE NON-INFRACTION VIS-À-VIS DES LOIS. Les produits IBM bénéficient d'une garantie, conformément aux conditions générales des contrats dans le cadre desquels ils sont fournis.

Le client est responsable du respect des lois et réglementations qui lui sont applicables. IBM n'assure aucun conseil juridique, et ne déclare ou ne garantit en aucune manière que ses services ou ses produits assurent la conformité du client avec une loi ou une réglementation quelconques.

Déclaration de bonnes pratiques de sécurité : La sécurité des systèmes informatiques implique la protection des systèmes et des informations via des mesures de prévention, de détection et de réponse aux accès non autorisés venant de l'intérieur ou de l'extérieur de votre entreprise. Les accès non autorisés peuvent entraîner la modification, la destruction, le détournement ou l'utilisation incorrecte d'informations et conduire à la détérioration ou à l'utilisation incorrecte de vos systèmes, notamment dans le but de lancer des attaques contre des tiers. Aucun système ou produit informatique ne doit être considéré comme entièrement sécurisé et aucun produit, service ou aucune mesure de sécurité n'est infaillible dans le cadre de la prévention des utilisations incorrectes ou des accès non autorisés. Les systèmes, produits et services IBM sont conçus pour être intégrés dans une solution de sécurité complète, qui comprend impérativement d'autres procédures opérationnelles et peut nécessiter une grande efficacité de la part d'autres systèmes, produits ou services. IBM NE GARANTIT D'AUCUNE MANIÈRE QUE LES SYSTÈMES, PRODUITS OU SERVICES SONT EXEMPTS DE TOUT ÉLÉMENT MALVEILLANT OU DE TOUTE ACTION ILLÉGALE PROVENANT DE TIERS OU QU'ILS PROTÈGERONT VOTRE ENTREPRISE CONTRE DE TELS ÉLÉMENTS.

- 1 IBM Institute for Business Value: IBM Global C-suite Study, 20e édition, « [Build Your Trust Advantage, Leadership in the era of data and AI everywhere](#) », novembre 2019.
- 2 Forrester Research, selon une étude menée par Forrester Consulting pour le compte d'IBM, août 2019.
- 3 Sangeet Paul Choudary, « [Platform Revolution: How Networked Markets Are Transforming the Economy—and How to Make Them Work for You](#) », 2016.
- 4 PwC, « [2019 AI Predictions, Six AI priorities you can't afford to ignore](#) », pwc.com, 2019.
- 5 Julianna Delua, « [AutoAI wins Alconics Intelligent Automation Award: Meet a key inventor](#) », ibmbigdatahub.com, septembre 2019.
- 6 Sajan Kuttappa, « [The rise of decision intelligence: AI that optimizes decision-making](#) », ibmbigdatahub.com, octobre 2019.
- 7 Paul Kilroy, « [Industry Accelerators for Cloud Pak for Data](#) », **ibm.com**, juillet 2019.
- 8 Carlo Appugliese, Paco Nathan et William S. Roberts, « [Agile AI: A Practical Guide to Building AI Applications and Teams](#) », O'Reilly eBooks, 2019.

