



keyrus
make data matter

TAB
Tech Advisory Board

Data matriochkas

IA, machine learning & deep learning

www.keyrus.com

Data matriochkas

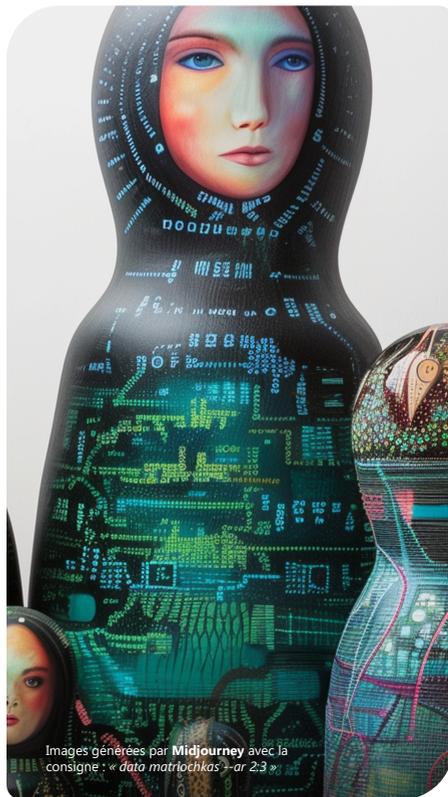
IA, machine learning & deep learning

Dans l'univers fascinant de la technologie, l'intelligence artificielle se présente comme une poupée russe, où chaque avancée cache en son sein une innovation encore plus spécialisée. La poupée russe de l'intelligence artificielle est une image qui illustre parfaitement cette idée de superposition et d'englobement, où des concepts comme le machine learning, le deep learning, et les modèles de langage à grande échelle cohabitent et s'imbriquent les uns dans les autres dans une hiérarchie subtile et captivante.

La plus grande matriochka incarne la data science, un domaine vaste et englobant, qui utilise des méthodes statistiques pour tirer des connaissances et des insights à partir de données brutes. Cette couche externe ou meta-poupée recouvre l'intelligence artificielle, qui à son tour, enveloppe le machine learning, lequel renferme les réseaux de neurones, incluant les techniques de deep learning, qui elles-mêmes, contiennent les LLM.

Et comment démêler cet enchevêtrement, nous direz-vous ? Tout simplement en posant les enjeux actuels et futurs de ces technologies, en définissant clairement chaque composante de notre poupée russe technologique, et en discutant des systèmes experts, ces algorithmes complexes et arbres de décision qui ont pavé la voie à l'IA d'aujourd'hui.

Par-delà cette exploration, nous n'insisterons jamais trop sur l'importance fondamentale de la qualité des données. Notre conviction restant qu'une approche réussie en IA et ML ne peut être déconnectée du réel et doit toujours commencer par une base solide : des données de qualité.



Images générées par Midjourney avec la consigne : « data matriochkas » - ar 2.3 »

C'est donc une plongée dans la data avec le scaphandrier à plusieurs couches qui compose l'IA que cet article vous invite à réaliser, en expliquant comment chaque niveau s'appuie sur le précédent pour créer des systèmes de plus en plus sophistiqués.

Last but not least, une lecture pour comprendre comment Keyrus accompagne ses clients, des novices aux plus avancés, dans la conduite de leurs projets IA, en soulignant l'importance d'une approche méthodologique rigoureuse et d'une spécialisation pointue pour répondre aux défis de l'industrialisation de l'IA. Da !

La poupée russe technologique

La métaphore de la poupée russe offre une illustration évocatrice pour comprendre la structure complexe et hiérarchisée de la technologie moderne, en particulier dans les domaines de l'intelligence artificielle (IA), du machine learning (ML), du deep learning (DL), et des modèles de langage à grande échelle (LLM). Chacune de ces technologies s'inscrit dans une strate plus large, englobant et étant englobée par d'autres, formant un ensemble cohérent et interdépendant.

Data science : la meta-poupée

La data science est un domaine interdisciplinaire qui s'appuie sur des méthodes statistiques, l'analyse de données, et l'apprentissage automatique pour extraire des connaissances et des insights à partir de données structurées et non structurées. Elle fusionne des domaines tels que les statistiques, l'informatique, l'ingénierie des données, et la visualisation des données pour analyser et interpréter des ensembles de données complexes.

L'objectif de la data science est de découvrir des patterns, de prédire des événements futurs, et de fournir des décisions basées sur des données dans divers contextes. Précisons en outre qu'elle englobe divers domaines, dont l'IA, offrant les fondements nécessaires pour comprendre et manipuler des volumes massifs de données et pour en tirer des conclusions significatives.

Dans le secteur de la fabrication ou du retail, par exemple, **la data science est utilisée pour optimiser les chaînes d'approvisionnement**, réduisant ainsi les coûts et améliorant l'efficacité. En analysant des données historiques et en temps réel concernant la demande des consommateurs, les délais de livraison, les conditions météorologiques, et d'autres variables, les entreprises peuvent prédire plus précisément la demande future.



Cela leur permet d'ajuster leur production, leur stock et leur logistique pour éviter les surstocks ou les ruptures de stock, garantissant ainsi que les produits sont disponibles quand et où les clients en ont besoin. Le parfait exemple désormais célèbre est celui d'une entreprise de vente au détail peut utiliser la data science pour anticiper les pics de demande saisonniers et planifier à l'avance son stock et ses ressources logistiques.

Les entreprises peuvent également utiliser la data science pour **analyser les sentiments des consommateurs** à partir de données collectées sur les réseaux sociaux, les avis en ligne, les forums, et d'autres plateformes numériques. En appliquant des techniques de traitement du langage naturel et d'analyse de texte, elles peuvent identifier les opinions positives, négatives, ou neutres des clients concernant leurs produits ou services.

Cette compréhension permet aux entreprises de détecter rapidement les problèmes ou les aspects négatifs perçus, d'ajuster leur stratégie de communication, de répondre aux préoccupations des clients, et de mettre en œuvre des améliorations produit basées sur les retours des utilisateurs. Par exemple, une entreprise de cosmétiques peut analyser les avis des clients sur ses produits pour identifier les ingrédients ou les caractéristiques les plus appréciés et ceux nécessitant des ajustements.

Ainsi, la data science permet aux entreprises et aux organisations de prendre des décisions éclairées et stratégiques en s'appuyant sur l'analyse de données. En transformant les données brutes en insights actionnables, la data science devient un outil puissant pour améliorer les opérations, la satisfaction client, et l'innovation produit.



Image générée par Midjourney avec la consigne : « AI-enhanced diagnostic systems serve as a doctor's digital assistant, sifting through mountains of patient data with lightning speed and pinpoint accuracy. By matching patterns in the data with an extensive library of medical conditions, these systems help doctors diagnose diseases faster and more accurately, revolutionizing patient care. »

En se fondant sur les insights et les connaissances extraites par la data science, nous franchissons le seuil vers l'intelligence artificielle, un domaine où ces données sont non seulement interprétées, mais aussi **utilisées pour simuler l'intelligence humaine et permettre aux machines de prendre des décisions autonomes.**

Intelligence artificielle : la fille aînée

L'intelligence artificielle représente un champ vaste et dynamique de la technologie qui cherche à créer des systèmes capables de simuler divers aspects de l'intelligence humaine. Ces aspects incluent l'apprentissage, le raisonnement, la perception, la compréhension du langage, et la capacité à planifier et à s'adapter à de nouvelles situations.

Au-delà de la simple automatisation de tâches répétitives, l'IA vise à fournir des solutions qui peuvent améliorer la prise de décision, optimiser les processus complexes et innover dans la manière dont nous interagissons avec le monde autour de nous.

En intégrant des techniques telles que la machine learning, qui permet aux systèmes d'apprendre et d'améliorer leurs performances sans être explicitement programmés pour chaque cas, et les systèmes experts, qui utilisent des ensembles de règles pour résoudre des problèmes spécifiques, l'IA a le potentiel de transformer radicalement les industries et notre quotidien.

Dans le domaine de la santé, l'IA est utilisée pour développer des systèmes de diagnostic assisté qui peuvent aider les médecins à identifier et à diagnostiquer les maladies avec une précision et une vitesse supérieures. En analysant des images médicales, comme les radiographies, les IRM ou les scans CT, les algorithmes d'IA peuvent détecter des signes précoces de maladies, telles que le cancer, souvent avant qu'ils ne soient visibles pour l'œil humain.

Cette capacité permet une intervention plus rapide et peut considérablement améliorer les chances de succès du traitement. Par exemple, des systèmes d'IA ont été entraînés pour identifier avec précision les tumeurs mammaires malignes, aidant ainsi les radiologues à réduire le taux de faux négatifs et de faux positifs.

((L'objectif de la data science est de **découvrir des patterns**, de **prédire des événements futurs**, et de **fournir des décisions basées sur des données** dans divers contextes.))



Image générée par Midjourney avec la consigne :
 « AI capable of understanding natural language and responding to customer questions, solving problems, and providing information about products or services 24 hours a day, 7 days a week --ar 3:8 --v 6.0. »

Les entreprises de divers secteurs utilisent des chatbots alimentés par l'IA pour fournir un service client instantané et personnalisé à grande échelle (sur ce sujet spécifiquement, nous vous recommandons la lecture de notre *Customer Care Playbook*). Ces chatbots sont capables de comprendre le langage naturel et de répondre aux questions des clients, de résoudre des problèmes et de fournir des informations sur les produits ou services 24 heures sur 24, 7 jours sur 7. L'avantage ici est double : améliorer l'expérience client en fournissant des réponses rapides et précises à leurs questions, et réduire les coûts pour les entreprises en automatisant les tâches de service client répétitives. Par exemple, une banque peut utiliser un chatbot IA pour aider ses clients à naviguer dans leurs options de compte, à signaler des cartes perdues ou volées, et même à fournir des conseils financiers personnalisés.

Ces exemples démontrent l'étendue et la profondeur de l'impact que l'IA peut avoir sur divers domaines, offrant des solutions innovantes qui améliorent à la fois l'efficacité opérationnelle et la qualité de vie.

Si nous continuons d'ouvrir nos différentes poupées russes, au cœur de l'intelligence artificielle se trouve le machine learning, une approche qui transcende la programmation traditionnelle en permettant aux systèmes d'apprendre et d'évoluer à partir des données elles-mêmes, sans instructions explicites pour chaque tâche spécifique.

Machine learning : au cœur de l'IA

Le machine learning constitue une avancée majeure dans le domaine de l'intelligence artificielle, marquant une évolution significative dans la manière dont les machines peuvent traiter l'information et apprendre. Contrairement aux programmes informatiques traditionnels, qui suivent des instructions précises pour effectuer des tâches, le ML repose sur des algorithmes capables d'apprendre de manière autonome à partir des données.

En ingérant de vastes ensembles de données et en identifiant des schémas ou des tendances au sein de ces données, les modèles de ML s'ajustent et améliorent leurs performances sans intervention humaine directe. Cette **capacité à apprendre et à s'adapter** rend le ML extrêmement puissant pour une gamme étendue d'applications, allant de la reconnaissance d'image à la prédiction de tendances.

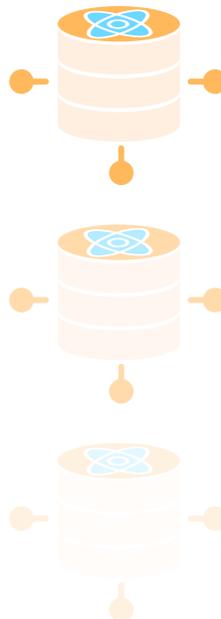
Les institutions financières emploient des modèles de machine learning pour identifier les transactions frauduleuses en temps réel. En analysant des motifs dans les transactions historiques, les algorithmes peuvent apprendre à détecter des anomalies ou des comportements suspects qui pourraient indiquer une fraude. Cette capacité à identifier rapidement les transactions potentiellement frauduleuses permet aux banques de bloquer les transactions douteuses avant qu'elles ne soient finalisées, protégeant ainsi les clients et réduisant les pertes financières. Par exemple, un modèle de ML peut reconnaître une tentative de transaction dans un pays étranger peu de temps après un achat dans le pays d'origine du détenteur de la carte, signalant ainsi une possible fraude.

Autre cas d'usage, **les entreprises du retail** utilisent le machine learning pour anticiper la demande des consommateurs pour différents produits.

En analysant des données historiques de vente, y compris les tendances saisonnières, les événements spéciaux, et même les conditions météorologiques, les modèles de ML peuvent prédire avec précision les fluctuations de la demande. Cette prévision permet aux détaillants de gérer leur inventaire de manière plus efficace, en s'assurant qu'ils disposent du stock adéquat pour répondre à la demande sans encourir de surplus coûteux. Par exemple, une chaîne de supermarchés peut utiliser le ML pour optimiser ses commandes de produits frais, réduisant ainsi le gaspillage alimentaire tout en maximisant les ventes.

Mais c'est le même cas d'usage que celui présenté pour illustrer la data science, me direz-vous ! C'est presque vrai... La différence fondamentale entre la data science et le machine learning, surtout dans le contexte de la prévision de la demande dans le secteur de la vente au détail, réside dans leur portée, leurs méthodologies et leurs objectifs finaux.

La data science peut impliquer la **collecte** et le **nettoyage** des données, l'**analyse** exploratoire pour comprendre les tendances et les schémas dans les données historiques de vente, et l'**utilisation de statistiques** pour identifier les facteurs influençant la demande.



Le but est d'obtenir une compréhension globale de la demande des consommateurs à partir des données disponibles, ce qui peut inclure l'identification de corrélations, la réalisation de tests d'hypothèses, et la visualisation de tendances pour aider à la prise de décision.

Le machine learning, quant à lui, est une sous-discipline de l'IA qui se concentre spécifiquement sur la construction d'algorithmes capables d'apprendre de et de faire des prédictions ou des décisions basées sur des données. Dans l'application à la prévision de la demande, le ML va au-delà de l'analyse statistique traditionnelle en utilisant des modèles complexes qui peuvent apprendre automatiquement les relations et les patterns dans les données historiques de vente sans être explicitement programmés pour chaque cas. Les modèles de ML peuvent ajuster leurs paramètres en fonction des données d'entrée pour améliorer la précision de leurs prédictions concernant la demande future.

Cette approche est particulièrement utile pour gérer la complexité et la volatilité des comportements de consommation, en permettant des prédictions dynamiques qui s'adaptent à de nouvelles informations.

En résumé, la data science fournit les bases nécessaires pour comprendre les données de vente au détail, en utilisant une combinaison d'analyse statistique, de visualisation, et d'exploration de données pour identifier les tendances et les facteurs influençant la demande. Le machine learning, lui, applique des modèles prédictifs spécifiques qui apprennent de ces données pour faire des prédictions précises sur la demande future, en s'adaptant automatiquement aux changements dans les tendances des données sans intervention humaine directe. Dit autrement, alors que la data science offre le cadre et les outils nécessaires pour analyser et interpréter les données, le machine learning apporte une capacité prédictive avancée, permettant aux entreprises de vente au détail d'anticiper et de répondre plus efficacement aux fluctuations de la demande.

Ces exemples illustrent la polyvalence et l'efficacité du machine learning dans la résolution de problèmes complexes et la prise de décision basée sur les données.

Le machine learning sert de fondation au deep learning, une technique plus avancée qui imite la manière dont le cerveau humain opère, en utilisant des réseaux de neurones complexes pour digérer et interpréter des ensembles de données volumineux, ouvrant la voie à des avancées sans précédent dans la reconnaissance de motifs et la compréhension du langage.

Deep learning : l'essence du ML

Le deep learning représente une révolution dans le domaine de l'intelligence artificielle et du machine learning, grâce à son utilisation de réseaux de neurones profonds. Ces réseaux, constitués de multiples couches de neurones artificiels, imitent la manière dont le cerveau humain traite l'information. Cette architecture permet au deep learning de capturer des niveaux d'abstraction de plus en plus complexes au fur et à mesure que l'information se propage à travers ses couches. Cette capacité unique à traiter et à analyser des volumes massifs de données avec une précision élevée rend le DL particulièrement adapté à des applications impliquant des données non structurées ou complexes, telles que les images, le son et le langage naturel.

Un domaine où le deep learning a eu un impact significatif est la reconnaissance faciale, utilisée tant dans les applications de sécurité que dans les systèmes d'identification personnelle. Les entreprises de sécurité utilisent le DL pour développer des systèmes capables d'identifier avec précision des individus dans des foules ou à des points d'accès, améliorant ainsi la sécurité des espaces publics et privés.

Par exemple, les aéroports de certains pays utilisent la reconnaissance faciale pour accélérer le processus d'identification des passagers, tout en renforçant les mesures de sécurité. Cette technologie permet une identification rapide et fiable, réduisant les files d'attente et améliorant l'expérience utilisateur, tout en s'assurant que seules les personnes autorisées accèdent à des zones sécurisées.



Images générées par **Ideogram** avec les consignes : « A futuristic airport scene with a high-tech facial recognition system in action. A large screen displays the face of a passenger, while the system's camera scans their face. The background features a modern airport terminal with numerous travelers, and airport staff monitor the process with attentive expressions. The overall atmosphere is efficient, secure, and technologically advanced.»

“ La hiérarchie de la data science, de l’IA, du ML, du DL et des LLM, semblable à celle des poupées russes, met en lumière la manière dont **chaque nouvelle couche de technologie s’appuie sur la précédente pour atteindre un niveau de sophistication** et d’application toujours plus élevé. ”

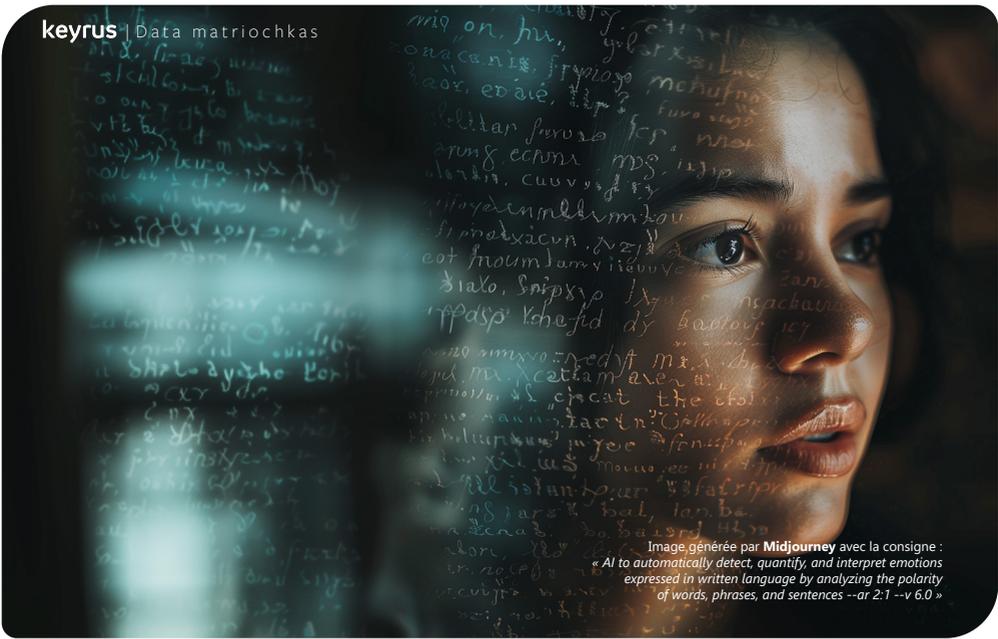


Image générée par Midjourney avec la consigne :
« AI to automatically detect, quantify, and interpret emotions expressed in written language by analyzing the polarity of words, phrases, and sentences --ar 2:1 --v 6.0 »

Le deep learning est également utilisé pour analyser les sentiments exprimés dans les textes publiés sur les réseaux sociaux, offrant aux entreprises des insights précieux sur les perceptions des consommateurs à l'égard de leurs marques ou produits. En traitant le langage naturel, les modèles de DL peuvent détecter les nuances sentimentales dans les avis des clients, qu'ils soient positifs, négatifs ou neutres. Cette information est essentielle pour les départements de marketing et de service clientèle, qui peuvent ainsi ajuster leurs stratégies en temps réel. Par exemple, une entreprise de cosmétiques peut utiliser l'analyse sentimentale pour surveiller les réactions aux lancements de nouveaux produits, identifier les points de satisfaction ou de mécontentement, et adapter ses campagnes de communication ou ses formulations de produits en conséquence.

En transcendant les limites des techniques traditionnelles de machine learning, le deep learning ouvre de nouvelles voies pour l'innovation et l'amélioration des processus dans une variété de secteurs, transformant non seulement la manière dont les entreprises opèrent, mais aussi leur approche en générale d'une multitude de problématiques.

Et il est possible de pousser les capacités du deep learning encore plus loin ! Nous arrivons à l'ère des modèles de langage à grande échelle, où les machines ne se contentent plus de reconnaître des patterns, mais commencent à comprendre et à générer du langage humain de manière cohérente, ouvrant ainsi de nouvelles dimensions dans l'interaction homme-machine.

Les modèles de langage à grande échelle : la quintessence du DL

Les modèles de langage à grande échelle (LLM pour *large language model*) représentent une avancée majeure dans le domaine du traitement automatique du langage naturel (NLP pour *natural language processing*), grâce à leur capacité à comprendre, générer et interagir en utilisant le langage humain de manière extrêmement sophistiquée. Basés sur des architectures de deep learning avancées, ces modèles sont entraînés sur des corpus de texte gigantesques, couvrant une vaste gamme de sujets, styles et contextes. Cette formation exhaustive leur permet de saisir les nuances du langage, y compris la grammaire, le ton, les idiomes et même l'humour, leur donnant ainsi la capacité de produire du texte qui peut être indiscernable de celui rédigé par un humain.

Les LLM peuvent accomplir une variété de tâches NLP, comme la traduction automatique, la synthèse de texte, la compréhension de questions et la génération de réponses pertinentes, transformant radicalement les possibilités d'interaction entre les humains et les machines.

Dans le domaine du service client, les LLM sont utilisés pour créer des assistants virtuels et des chatbots capables de fournir des réponses précises, contextuelles et naturelles aux questions des clients. Ces systèmes avancés peuvent comprendre les requêtes formulées en langage naturel, les analyser pour saisir l'intention de l'utilisateur, et générer des réponses informatives et pertinentes. Par exemple, une banque peut déployer un assistant virtuel alimenté par un LLM pour aider ses clients à naviguer dans ses services, à résoudre des problèmes courants ou à obtenir des conseils financiers personnalisés, améliorant ainsi l'expérience client tout en réduisant la charge sur son service d'assistance humain.

Alors que l'IA inclut diverses technologies pour développer des assistants et chatbots, comme l'analyse de données, le machine learning et la compréhension du langage naturel, les LLM fournissent la base technologique pour que ces assistants et chatbots comprennent et génèrent du langage de manière cohérente et contextuelle, rendant les interactions plus naturelles et efficaces.

Les LLM trouvent également leur application dans la génération automatique de contenu, où ils peuvent produire des textes de haute qualité sur une variété de sujets. Les entreprises de médias et de contenu utilisent ces modèles pour générer des articles de nouvelles, des résumés de contenu, ou même des scripts pour des vidéos et des podcasts. Par exemple, une agence de marketing numérique peut utiliser un LLM pour créer des descriptions de produits engageantes pour les catalogues en ligne de ses clients, des articles de blog sur des sujets spécifiques, ou du contenu pour les réseaux sociaux, réduisant significativement le temps et les ressources nécessaires à la création de contenu original et pertinent.

Ces exemples illustrent l'impact transformateur des LLM dans le traitement du langage naturel et la génération de texte, offrant des outils puissants pour améliorer l'interaction homme-machine et automatiser la création de contenu.

La hiérarchie de la data science, de l'IA, du ML, du DL et des LLM, semblable à celle des poupées russes, met en lumière la manière dont chaque nouvelle couche de technologie s'appuie sur la précédente pour atteindre un niveau de sophistication et d'application toujours plus élevé. Cette structure non seulement illustre l'évolution des capacités technologiques, mais souligne également l'importance de la compréhension et de l'expertise à chaque niveau pour exploiter pleinement le potentiel de l'IA moderne.



Image générée par Midjourney
avec la consigne : «customer service
enhanced with AI --ar 4:1.»

La donnée avant tout

Au pays de l'intelligence artificielle et des machine learning et deep learning, la qualité et la pertinence des données sont primordiales. Avant de plonger dans les complexités algorithmiques ou les architectures de réseau sophistiquées, il est essentiel de reconnaître que **le fondement de toute réussite en IA réside dans la donnée elle-même**. Commencer par la donnée est non seulement stratégique mais fondamental pour le succès des projets d'IA, ML, DL...

L'importance de la qualité des données ne saurait être sous-estimée. Les modèles sont conçus pour apprendre à partir des données fournies, ce qui signifie que la **précision**, la **cohérence** et la **complétude** de ces données jouent un rôle critique dans la détermination de l'efficacité et de la fiabilité des résultats obtenus. Des données de haute qualité sont essentielles pour garantir que les modèles puissent identifier correctement les patterns, effectuer des prédictions précises et fournir des insights valables.

- **La précision des données** fait référence à l'exactitude des informations collectées. Des données imprécises peuvent être le résultat d'erreurs de mesure, de saisie incorrecte des données, ou de sources d'information peu fiables. Dans les modèles d'IA et de ML, même de petites inexactitudes peuvent entraîner des erreurs significatives dans les prédictions ou les classifications. Assurer la précision des données dès le départ minimise le risque d'erreurs propagées à travers le modèle, garantissant ainsi que les décisions basées sur ces prédictions sont fondées sur des informations fiables.
- **La cohérence des données** se réfère à l'uniformité des formats de données, des valeurs et des mesures à travers l'ensemble de données. Des incohérences, telles que l'utilisation de différents formats de date ou d'unités de mesure, peuvent brouiller les patterns que les modèles d'IA tentent d'apprendre, réduisant leur capacité à généraliser à partir des données d'entraînement aux données inédites.

La normalisation des données avant leur utilisation dans des modèles d'IA aide à garantir que les algorithmes travaillent avec des informations homogènes, améliorant ainsi la qualité des insights générés.

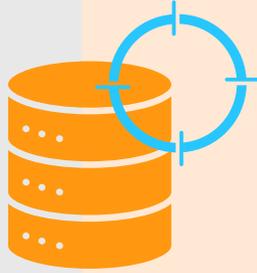
- **La complétude** fait référence à l'absence de valeurs manquantes dans un ensemble de données. Les données incomplètes peuvent être particulièrement problématiques pour les modèles d'IA et de ML, car elles peuvent induire des biais dans l'apprentissage ou obliger le modèle à faire des suppositions sur les données manquantes, ce qui peut entraîner des prédictions inexactes. Des techniques de gestion des valeurs manquantes, telles que l'imputation ou l'utilisation de modèles capables de gérer l'absence de données, sont indispensables pour maintenir l'intégrité des résultats.

La proverbiale maxime *garbage in, garbage out* n'a jamais été aussi pertinente : sans une base de données solide et fiable, même les algorithmes les plus avancés ne peuvent produire des résultats fiables.

Un modus operandi indépassable

On l'aura compris, la qualité et la pertinence des données utilisées pour entraîner ces systèmes sont fondamentales, car elles influencent directement la précision et l'efficacité des modèles générés. Et cette qualité repose sur quelques piliers et un processus de production des données, indépassables.

La première étape dans la préparation des données consiste à **identifier et à collecter les bonnes sources de données**. Cela nécessite une compréhension claire des objectifs du projet, ainsi que la connaissance fine de son secteur d'activité. Il faut sélectionner les sources de données les plus pertinentes, qu'elles soient internes à l'organisation ou externes, et comprendre les avantages et les limites de ces différentes sources.



Assurer la **précision** des données dès le départ minimise le risque d'erreurs propagées à travers le modèle.



La **normalisation** des données aide à garantir que les algorithmes travaillent avec des informations homogènes.



Des techniques de gestion des valeurs manquantes, telles que l'**imputation**, sont indispensables pour maintenir l'intégrité des résultats.

Images générées par Lexica aperture 3.5, ideogram et Midjourney avec la consigne : « Imagine a set of Russian dolls, each nested within the other, representing the intricate layers of the data science world. The largest doll, vibrantly painted with complex equations and datasets, symbolizes the vast field of Data Science itself. Inside it, you find a slightly smaller doll, adorned with intricate circuits and binary codes, representing Artificial Intelligence - a subset of data science that focuses on creating machines that mimic human intelligence. As you open the AI doll, you discover another doll, this one decorated with decision trees, neural networks, and various algorithms. This doll represents Machine Learning, a branch of AI that enables machines to learn from data and improve their performance without being explicitly programmed. Delving deeper, you find a smaller, more intricate doll, etched with layers of interconnected nodes resembling a web. This doll symbolizes Deep Learning, a subset of Machine Learning that focuses on training artificial neural networks to make decisions based on complex, abstract patterns. Finally, at the heart of all these layers, you find the smallest doll, inscribed with lines of text and dialogue, representing Large Language Models. These models, like the doll, are the smallest yet most intricate part of the data science world, capable of understanding and generating human-like text based on the patterns they've learned from vast amounts of data. »

Une fois les données collectées, elles doivent être nettoyées et préparées. Cette étape inclut la correction des erreurs, la gestion des valeurs manquantes, et la suppression des doublons. **Le nettoyage des données** est vital pour garantir la fiabilité des modèles d'IA et de ML. Les experts en data science jouent ici un rôle crucial en appliquant des techniques avancées de nettoyage et en veillant à ce que les données soient de la plus haute qualité possible.

L'intégration des données est une autre étape importante, surtout dans les environnements où les données proviennent de multiples sources hétérogènes. Elle implique de combiner ces données de manière cohérente, en s'assurant que les informations sont alignées et peuvent être utilisées de manière unifiée. Cette phase peut nécessiter l'utilisation d'outils d'intégration de données ou le développement de processus personnalisés pour assurer une intégration fluide.

L'enrichissement des données consiste à ajouter des informations supplémentaires aux données existantes pour améliorer leur utilité dans les modèles d'IA et de ML. Cela peut impliquer l'ajout de données contextuelles ou l'application de techniques d'ingénierie de caractéristiques pour créer de nouveaux attributs à partir des données existantes. Les experts en data science peuvent fournir des conseils précieux sur les meilleures méthodes d'enrichissement pour maximiser la valeur des données.

Enfin, **une stratégie de gestion de données** efficace est essentielle pour maintenir la qualité et la pertinence des données au fil du temps. Cela inclut la mise en place de processus pour la mise à jour régulière des données, la sécurisation des données sensibles, et la conformité avec les réglementations en vigueur.



Des poupées au monde réel

L'adoption d'une approche réaliste et connectée au monde réel dans la collecte et l'analyse des données est fondamentale pour le succès des initiatives IA et ML. Cette approche ne se contente pas de traiter les données en tant que telles ; elle les considère comme le reflet d'une réalité complexe, imprégnée de nuances commerciales, sociales et économiques. En reconnaissant et en intégrant cette complexité dans les modèles d'IA et de ML, les organisations peuvent s'assurer que leurs solutions ne sont pas seulement des exploits techniques, mais des outils réellement adaptés aux défis qu'elles cherchent à surmonter.

La première étape vers une approche réaliste est de **développer une compréhension profonde du contexte** dans lequel les données existent.

Cela signifie aller au-delà des chiffres pour saisir les dynamiques commerciales, les comportements des consommateurs, les tendances du marché et les facteurs socio-économiques qui influencent les données.

Par exemple, dans le secteur de la vente au détail, comprendre les variations saisonnières, les événements spéciaux et les préférences changeantes des consommateurs peut aider à affiner les modèles de prévision de la demande, rendant ces prévisions plus précises et, par conséquent, plus utiles pour la planification des stocks. Dit autrement, l'humain doit apporter sa vision et sa perception de l'environnement pour guider les algorithmes. La part d'intuition, de ressenti et d'émotion doit nécessairement se combiner à la partie plus scientifique. **Il faut tendre vers une fusion de la science et de l'art.**



Une approche réaliste implique également d'**aligner étroitement les projets d'IA et de ML sur les objectifs commerciaux de l'organisation.**

Cela nécessite une collaboration étroite entre les data scientists, les analystes de données (l'ensemble des **athlètes de la data**) et les parties prenantes commerciales pour s'assurer que les modèles développés répondent aux besoins réels de l'entreprise et contribuent de manière mesurable à ses objectifs stratégiques. Par exemple, un modèle conçu pour améliorer l'efficacité opérationnelle devrait être mesuré non seulement en termes de précision technique, mais aussi en termes d'impact sur les coûts, le temps de traitement ou la satisfaction client.

Reconnaître que les données et leur contexte peuvent changer avec le temps est également indispensable. Une approche réaliste prévoit la flexibilité et l'adaptabilité des modèles d'IA et de ML, permettant une mise à jour facile à mesure que de nouvelles informations deviennent disponibles ou que les conditions du marché évoluent. Cela garantit la pertinence continue des solutions d'IA et de ML et leur capacité à fournir des insights précieux, même dans un environnement changeant.

En fin de compte, l'adoption d'une approche réaliste et connectée au monde réel garantit que les solutions d'IA et de ML sont non seulement techniquement avancées, mais surtout pertinentes et directement applicables aux problèmes spécifiques qu'elles sont destinées à résoudre. Cela augmente considérablement les chances que ces solutions soient adoptées, intégrées et valorisées au sein de l'organisation, transformant ainsi les données en un véritable avantage concurrentiel.

Une approche réaliste dans la collecte et l'analyse des données inscrit les projets d'IA et de ML dans une perspective plus large, assurant que ces technologies ne sont pas seulement des exploits d'ingénierie, mais des outils pragmatiques conçus pour répondre aux défis réels de l'organisation.

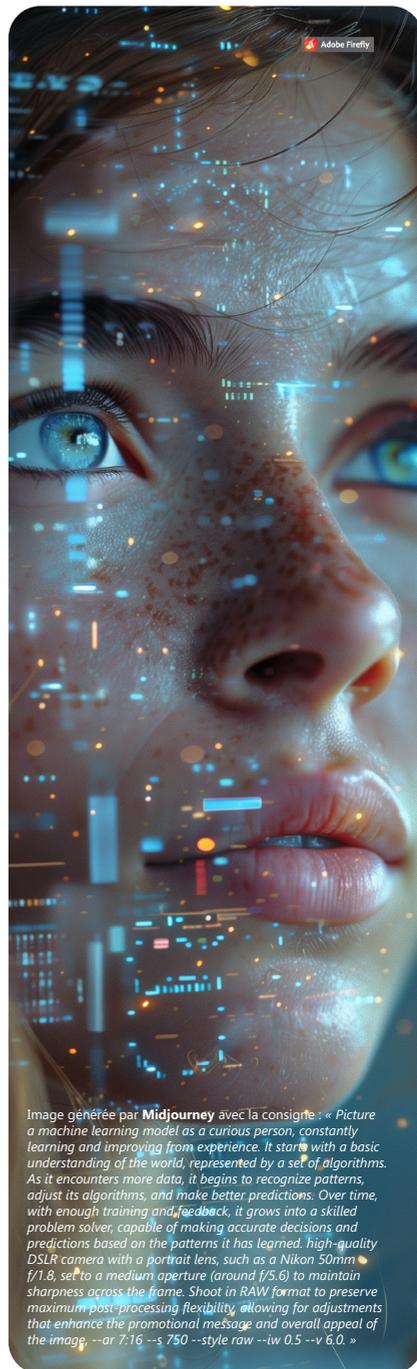


Image générée par **Midjourney** avec la consigne : « *Picture a machine learning model as a curious person, constantly learning and improving from experience. It starts with a basic understanding of the world, represented by a set of algorithms. As it encounters more data, it begins to recognize patterns, adjust its algorithms, and make better predictions. Over time, with enough training and feedback, it grows into a skilled problem solver, capable of making accurate decisions and predictions based on the patterns it has learned. high-quality DSLR camera with a portrait lens, such as a Nikon 50mm f/1.8, set to a medium aperture (around f/5.6) to maintain sharpness across the frame. Shoot in RAW format to preserve maximum post-processing flexibility, allowing for adjustments that enhance the promotional message and overall appeal of the image. --ar 7:16 --s 750 --style raw --iw 0.5 --v 6.0.* »

Image générée par Midjourney avec la consigne : « a data expert graces the scene with a cheerful hand gesture while showcasing a methodology diagram thoughtfully designed, inviting creative data application ideas, office environment background Employ a high-quality DSLR camera with a portrait lens, such as a Nikon 50mm f/1.8, set to a medium aperture (around f/5.6) to maintain sharpness across the frame. Shoot in RAW format to preserve maximum post-processing flexibility, allowing for adjustments that enhance the promotional message and overall appeal of the image. --ar 16:9 --s 750 --style raw --iw 0.5 »

Accompagnement et industrialisation des projets

La trajectoire de l'intelligence artificielle se dirige vers une ère marquée par la spécialisation et l'adoption d'applications ciblées. Cette évolution souligne un changement significatif dans la manière dont nous envisageons, concevons et déployons les technologies d'IA, avec un accent croissant sur la création de solutions hautement spécialisées, adaptées aux besoins spécifiques des secteurs et des entreprises. Keyrus, en tant qu'hyper spécialiste, s'inscrit au cœur de cette dynamique, offrant une expertise pointue et des solutions sur mesure qui répondent aux défis uniques de chaque client.

Keyrus propose en effet un accompagnement personnalisé à ses clients dans la mise en place de solutions d'intelligence artificielle. Cette méthode s'appuie sur une compréhension approfondie des différents degrés de familiarité avec l'IA au sein de sa clientèle. En ajustant finement ses interventions aux besoins uniques de chaque client, qu'ils débutent dans l'IA ou qu'ils disposent déjà d'une expertise avancée, Keyrus optimise l'efficacité des solutions d'IA déployées.

Pour les clients novices, l'objectif est de démystifier l'IA et de les guider à travers le processus d'identification des opportunités où l'IA pourrait apporter une valeur ajoutée significative. Cela implique souvent des ateliers de brainstorming pour définir des cas d'usage pertinents et des sessions de formation pour familiariser les clients avec les principes de base de l'IA et de la data science.

L'accent est mis sur la création d'une feuille de route claire pour l'implémentation de solutions d'IA qui peuvent transformer leurs opérations commerciales, améliorer l'efficacité et favoriser l'innovation.

Pour les clients avancés, qui possèdent déjà une compréhension solide de l'IA et ont identifié des projets spécifiques, l'accompagnement se concentre sur le peaufinage technique et stratégique. Cela peut inclure la sélection de la technique la plus appropriée – qu'il s'agisse de machine learning ou de deep learning – en fonction du cas d'usage spécifique. L'expertise de Keyrus dans divers domaines de l'IA permet d'orienter ces clients vers les solutions les plus efficaces, en s'assurant que les projets sont alignés avec les objectifs commerciaux et optimisés pour la performance.

Keyrus adopte par ailleurs une approche by design qui met l'accent sur la nécessité de concevoir des solutions avec une intention claire et une finalité précise dès le départ. Il propose surtout une vision pragmatique de l'industrialisation des projets d'IA, visant à créer des solutions qui sont non seulement innovantes mais aussi durables et économiquement viables. L'objectif est d'éviter que la mise en production et les ajustements des systèmes d'IA ne deviennent prohibitifs en termes de coûts, rivalisant avec les investissements initiaux des projets pilotes. Pour cela, Keyrus met en place un cadre méthodologique dès le départ, intégrant des pratiques de gestion de données avancées et des processus d'optimisation continue.



Image générée par **Lexica Aperture v4** avec la consigne :
« *Agraphic 70s poster, portrait of a data scientist, data miner,
data analyst, data strategist --ar 3:2, volumetric lighting.* »

Une stratégie d'entreprise réussie

En embrassant le voyage à travers les différentes strates de l'intelligence artificielle, du machine learning, du deep learning, jusqu'aux modèles de langage à grande échelle, nous avons mis en lumière l'importance de la qualité des données, la nécessité d'une approche individuelle et personnalisée, et l'impératif d'industrialisation efficace des projets d'IA.

La trajectoire de l'IA, désormais orientée vers la spécialisation et les applications ciblées, révèle une ère où la précision, l'efficacité, et l'innovation sont au premier plan. Et Keyrus, en tant qu'hyper spécialiste, incarne cette évolution, guidant les entreprises à travers le labyrinthe complexe de la data science vers la réalisation de leur potentiel numérique.

Ainsi, tout comme les poupées russes s'emboîtent les unes dans les autres, les différentes composantes de l'IA se complètent et s'articulent ensemble, soutenues par une fondation solide de données de qualité, et se révèlent être les pièces maîtresses d'une stratégie d'entreprise réussie.

L'avenir de l'IA et du ML se dévoile comme une série de *data matriochkas*, où chaque niveau de complexité débloque une valeur ajoutée supplémentaire, révélant le potentiel caché dans les données au cœur de l'innovation technologique.

**Article co-écrit par keyrus,
Chat-GPT⁴, Claude, Mistral et Gemini**

Vous avez trouvé cette lecture utile ?

Vous souhaitez sûrement lire aussi :

Innovation

L'ère du low-code / no-code

L'eBook *Innovation* explore comment le low-code/no-code révolutionne le développement d'applications, rendant la technologie accessible à tous. Il met en lumière les avantages considérables, tels que l'accélération de la transformation numérique et la réduction des coûts, tout en soulignant l'importance d'une stratégie réfléchie pour surmonter les défis associés. C'est un guide essentiel pour toute organisation aspirant à devenir plus adaptable et compétitive dans le paysage numérique actuel.

Quelles sont les 3 idées principales ?

1. Démocratisation du développement d'applications

: le low-code/no-code rend le développement accessible à un public plus large, y compris ceux sans compétences techniques approfondies, facilitant ainsi l'innovation et l'accélération de la transformation numérique.

2. Avantages pour les entreprises

: ces technologies offrent des avantages significatifs tels que la réduction des délais et des coûts de développement, améliorant l'agilité et la capacité d'adaptation des entreprises aux évolutions du marché.

3. Défis et considérations stratégiques

: malgré les avantages, l'adoption du low-code/no-code présente des défis comme la sélection de la bonne plateforme et la gestion des risques de sécurité. Une approche stratégique est essentielle pour maximiser les bénéfices tout en minimisant les risques.



keyrus
mobile data matter

Innovation

L'ère du low-code/no-code

www.keyrus.com



SCAN ME



Acteur international du conseil et des technologies, Keyrus a pour mission de donner du sens aux données, en révélant toute leur portée, notamment sous un angle humain.

Parce que ce ne sont pas tant les données elles-mêmes qui importent, mais les opportunités que nous pouvons développer en les apprivoisant vraiment, nous nous efforçons constamment de comprendre les objectifs que nos clients souhaitent atteindre. Nous explorons et mesurons les comportements, nous les comprenons et les traduisons en un résultat concret. Nous donnons un sens aux réalités que les données portent afin d'aider nos clients à prendre des décisions plus efficaces.

Les données, qu'elles soient grandes, petites, humaines, complexes, historiques ou prospectives, n'ont de sens que lorsqu'elles sont utilisées pour développer les expériences, affiner la compréhension du quotidien et prendre les meilleures décisions.

Notre proposition de valeur est fondée sur cinq grands groupes de services, chacun comprenant des offres multiples :

- **Automatisation et intelligence artificielle** : nous fournissons à nos clients les moyens d'améliorer leur productivité et leur précision sur l'ensemble de leurs processus, afin de se concentrer sur le travail à plus forte valeur ajoutée.
- **Expérience numérique centrée sur l'humain** : la relation avec les clients et l'engagement des collaborateurs constituent deux des plus grands contributeurs au succès global des entreprises. Nous aidons les entreprises à imaginer et à créer des expériences numériques multimodales et fluides pour atteindre leurs objectifs.
- **Mise en œuvre des données et des analyses** : les données sont une clé incontestable du succès pour les entreprises. Lorsqu'elles sont utilisées intelligemment, elles ouvrent des opportunités uniques pour faire face aux défis actuels et futurs. Nous permettons aux organisations de déployer tout le potentiel de leurs données : nous mettons la science des données au profit du développement de l'entreprise.
- **Cloud et sécurité** : le Cloud et les plateformes numériques ont le potentiel de révolutionner la façon dont les données sont transformées en valeur, tout en portant l'extensibilité et la flexibilité à un niveau supérieur. Nous sécurisons l'ensemble de vos données et veillons à ce qu'elles soient protégées et confidentielles.
- **Transformation et innovation** : pour prospérer dans l'écosystème actuel, chaque entreprise doit non seulement accélérer sa transformation numérique, mais aussi acquérir des compétences pour stimuler son adaptabilité, sa résilience et sa compétitivité. Nous aidons nos clients à se transformer avec succès pour développer un meilleur futur.

S'appuyant sur l'expérience cumulée de plus de 3 500 collaborateurs et présent dans 27 pays sur 4 continents, Keyrus est l'un des principaux experts internationaux en matière de données, de conseil et de technologie.

Pour en savoir plus : www.keyrus.fr

Jean-Philippe CLAIR

Directeur Marketing, Communication & Expérience client
jean-philippe.clair@keyrus.com