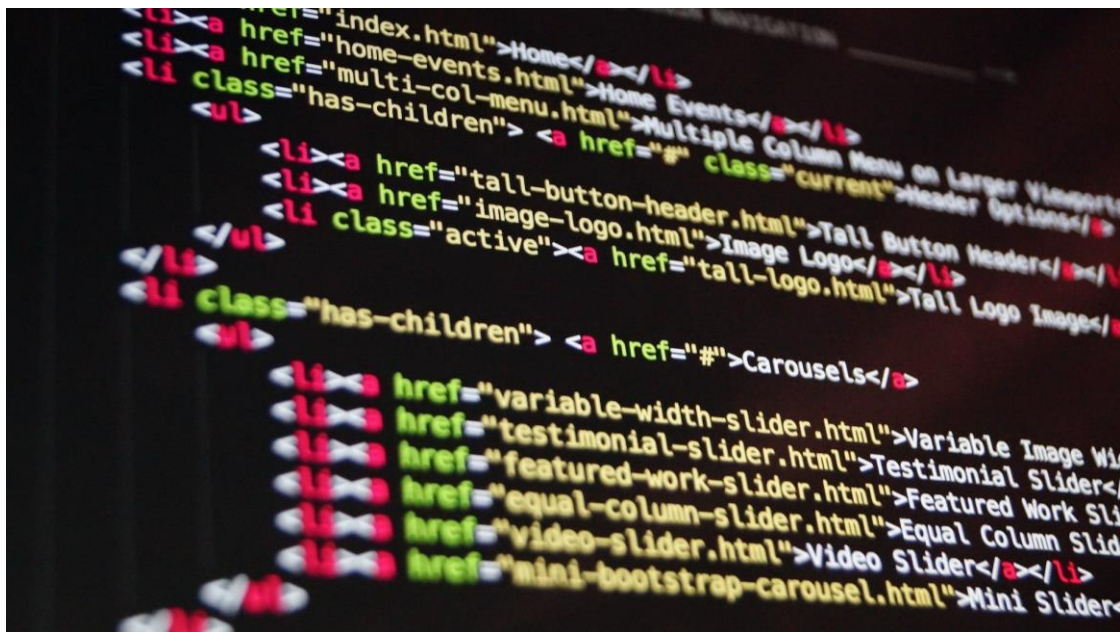


MACHINE LEARNING

BTS SIO option SLAM 2021-2023



Enzo NZUGUEM

TABLE DES MATIERES

| | |
|--------------------------------|----------|
| I CYCLE DE VEILLE | 1 |
|--------------------------------|----------|

| | |
|----------------------------------|----------|
| II MACHINE LEARNING | 2 |
|----------------------------------|----------|

A) Ciblage

B) Collecte

C) Traitement et Analyse

D) Diffusion

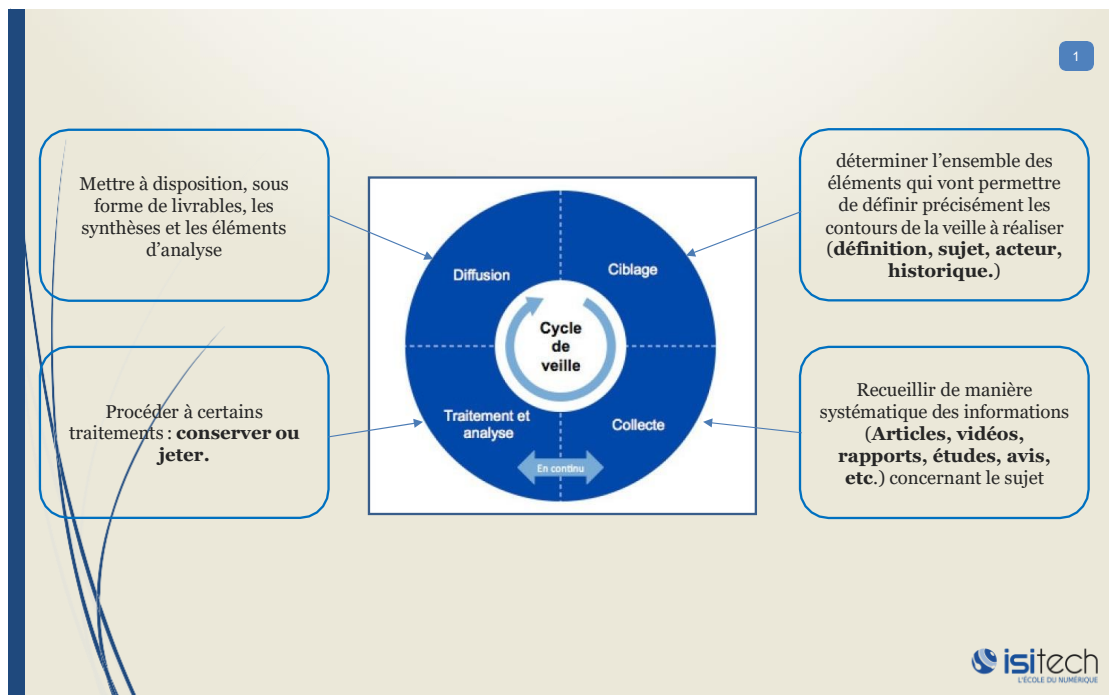
| | |
|---|----------|
| III LE MACHINE LEARNING DANS LE FUTUR..... | 1 |
|---|----------|

| | |
|----------------------------|----------|
| IV CONCLUSION | 1 |
|----------------------------|----------|

I- Cycle de veille

La mise en place et la réalisation de la veille technologique, reposera sur l'utilisation d'un processus de cycle de veille basé sur quatre étapes principales.

- Ciblage
- Collecte
- Traitement et Analyse
- Diffusion



II- Machine Learning

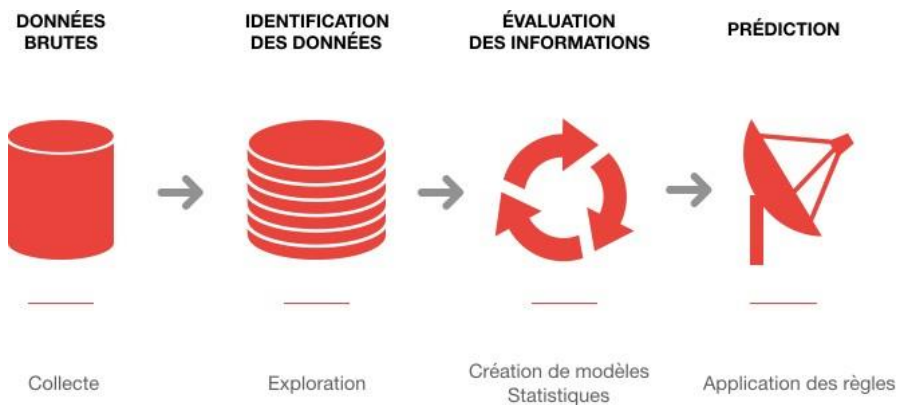
A) Ciblage

Définition Machine Learning :

L'apprentissage automatique (en anglais : machine learning) est un champ d'étude de l'intelligence artificielle qui se fonde sur des approches mathématiques et statistiques pour donner aux ordinateurs la capacité d'« apprendre » à partir de données, c'est-à-dire d'améliorer leurs performances à résoudre des tâches sans être explicitement programmés pour chacune. Plus largement, il concerne la conception, l'analyse, l'optimisation, le développement et l'implémentation de telles méthodes.

- Comment fonctionne le ML ?

Le machine learning fonctionne de tel sorte qu'on collecte dans un premier temps des données brutes, comme des images ou des sons. Ensuite il faut classer ces données par résultats que l'on veut obtenir, dans des dossiers par exemple. Dans un troisième temps on peut s'aider de logiciel comme CoreML, pour Apple, et TensorFlow pour les autres plateformes, ou bien de coder soit même sont programmes d'apprentissage de façon à soit le faire apprendre une fois soit le faire apprendre au fur et à mesure de son utilisation. Une fois le fichier entraîné il faut le mettre en place, dans une application mobile par exemple, et on peut faire afficher le résultat ainsi que le pourcentage de réussite de notre recherche.



- A quoi sert le machine learning ?

Le machine learning peut permettre de reconnaître des images et de les classer comme par exemple la reconnaissance de panneau sur les voitures récentes. Il peut aussi permettre de reconnaître des sons comme le fait l'application Shazam qui reconnaît les musiques ou bien Google Assistant ou Siri qui reconnaît la voix de l'utilisateur et lui permet d'effectuer des actions en fonction de ses demandes.

- Dans quoi est-il utilisé ?

L'apprentissage automatique est utilisé dans un large spectre d'applications pour doter des ordinateurs ou des machines de capacité d'analyser des données d'entrée comme : perception de leur environnement (vision, Reconnaissance de formes tels des visages, schémas, segmentation d'image, langages naturels, caractères dactylographiés ou manuscrits ; moteurs de recherche, analyse et indexation d'images et de vidéo, en particulier pour la recherche d'image par le contenu ; aide aux diagnostics, médical notamment, bio-informatique, chimio-informatique ; interfaces cerveau-machine ; détection de fraudes à la carte de crédit, cybersécurité, analyse financière, dont analyse du marché boursier ; classification des séquences d'ADN ; jeu ; génie logiciel ; adaptation de sites Web ; robotique (locomotion de robots, etc.) ; analyse prédictive dans de nombreux domaines (financière, médicale, juridique, judiciaire).

Par exemple un système d'apprentissage automatique peut permettre à un robot ayant la capacité de bouger ses membres, mais ne sachant

initialement rien de la coordination des mouvements permettant la marche, d'apprendre à marcher. Le robot commencera par effectuer des mouvements aléatoires, puis, en sélectionnant et privilégiant les mouvements lui permettant d'avancer, mettra peu à peu en place une marche de plus en plus efficace.

La reconnaissance de caractères manuscrits est une tâche complexe car deux caractères similaires ne sont jamais exactement identiques. Il existe des systèmes d'apprentissage automatique qui apprennent à reconnaître des caractères en observant des « exemples », c'est-à-dire des caractères connus. Un des premiers systèmes de ce type est celui de reconnaissance des codes postaux US manuscrits issu des travaux de recherche de Yann Le Cun, un des pionniers du domaine, et ceux utilisés pour la reconnaissance d'écriture ou OCR.

- **Quels sont les acteurs principaux ?**

Arthur Samuel, informaticien américain pionnier dans le secteur de l'intelligence artificielle, est le premier à faire usage de l'expression *machine learning* (en français, « apprentissage automatique ») en 1959 à la suite de la création de son programme pour IBM en 1952. Le programme jouait au Jeu de Dames et s'améliorait en jouant. À terme, il parvint à battre le 4ème meilleur joueur des États-Unis.

Une avancée majeure dans le secteur de l'intelligence machine est le succès de l'ordinateur développé par IBM, Deep Blue, qui est le premier à vaincre le champion mondial d'échecs Garry Kasparov en 1997. Le projet Deep Blue en inspirera nombre d'autres dans le cadre de l'intelligence artificielle, particulièrement un autre grand défi : IBM Watson, l'ordinateur dont le but est de gagner au jeu Jeopardy!. Ce but est atteint en 2011, quand Watson gagne à Jeopardy! en répondant aux questions par traitement de langage naturel.

Durant les années suivantes, les applications de l'apprentissage automatique médiatisées se succèdent bien plus rapidement qu'auparavant.

En 2012, un réseau neuronal développé par Google parvient à reconnaître des visages humains ainsi que des chats dans des vidéos YouTube.

En 2014, 64 ans après la prédiction d'Alan Turing, le dialogueur Eugene Goostman est le premier à réussir le test de Turing en parvenant à convaincre 33 % des juges humains au bout de cinq minutes de conversation qu'il est non pas un ordinateur, mais un garçon ukrainien de 13 ans.

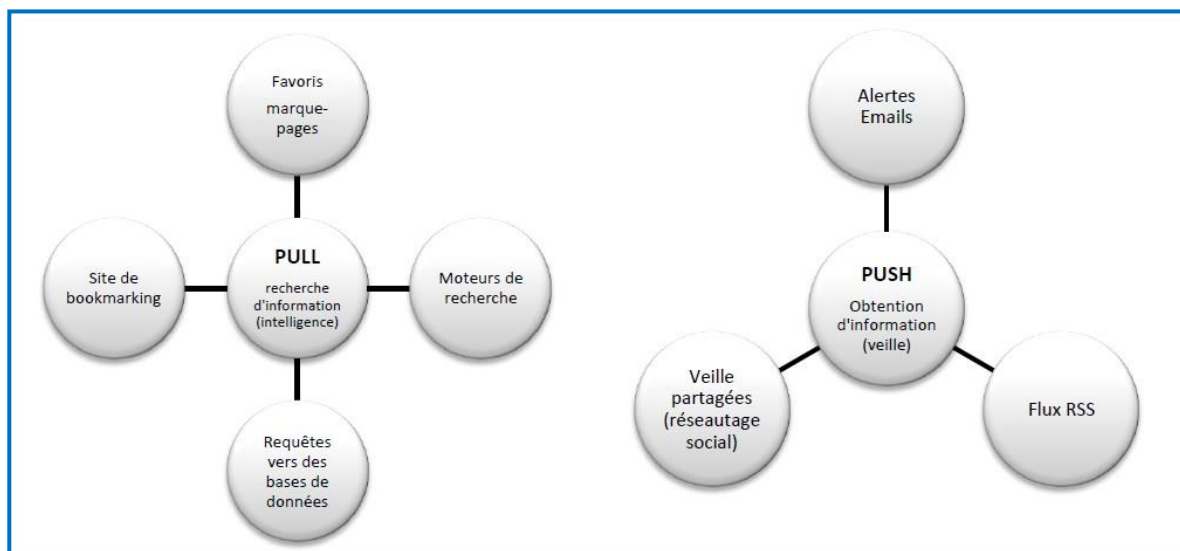
En 2015, une nouvelle étape importante est atteinte lorsque l'ordinateur « AlphaGo » de Google gagne contre un des meilleurs joueurs au jeu de Go, jeu de plateau considéré comme le plus dur du monde.

En 2016, un système d'intelligence artificielle à base d'apprentissage automatique nommé LipNet parvient à lire sur les lèvres avec un grand taux de succès.

B) Collecte

Tous d'abord la méthode « Pull » est une méthode qui consiste à ce que l'internaute aille chercher lui-même les informations dont il a besoin.

Ensuite la méthode « Push », elle, est une méthode qui consiste à faire venir l'information jusqu'à l'internaute de manière automatisée à l'aide de Flux RSS ou de Google Actualités par exemple.



C) Traitement et Analyse

Beaucoup d'informations sur internet parle du Machine Learning il faut donc les lire puis les trier pour ne sélectionner que celles qui sont intéressantes et utiles pour ma veille technologique.

Pour cela je vais donc m'orienter sur des sites spécialisés en Intelligence Artificielle et plus particulièrement en Machine Learning pour avoir un maximum d'informations et qui de plus sont à jour puisqu'elles proviennent de grands sites internationaux à la pointe du Machine Learning.

D) Diffusion

Les différents types de Machine Learning

On distingue différents types d'algorithmes Machine Learning. Généralement, ils peuvent être répartis en deux catégories : supervisés et non supervisés.

Dans le cas de l'apprentissage supervisé, les données utilisées pour l'entraînement sont déjà « étiquetées ». Par conséquent, le modèle de Machine Learning sait déjà ce qu'elle doit chercher (motif, élément...) dans ces données. À la fin de l'apprentissage, le modèle ainsi entraîné sera capable de retrouver les mêmes éléments sur des données non étiquetées.

Parmi les algorithmes supervisés, on distingue les algorithmes de classification (prédictions non-numériques) et les algorithmes de régression (prédictions numérique). En fonction du problème à résoudre, on utilisera l'un de ces deux archétypes.

L'apprentissage non supervisé, au contraire, consiste à entraîner le modèle sur des données sans étiquettes. La machine parcourt les données sans aucun indice, et tente d'y découvrir des motifs ou des tendances

récurrents. Cette approche est couramment utilisée dans certains domaines, comme la cybersécurité.

Parmi les modèles non-supervisés, on distingue les algorithmes de clustering (pour trouver des groupes d'objets similaires), d'association (pour trouver des liens entre des objets) et de réduction dimensionnelle (pour choisir ou extraire des caractéristiques).

Une troisième approche est celle de l'apprentissage par renforcement. Dans ce cas de figure, l'algorithme apprend en essayant encore et encore d'atteindre un objectif précis. Il pourra essayer toutes sortes de techniques pour y parvenir. Le modèle est récompensé s'il s'approche du but, ou pénalisé s'il échoue.

En tentant d'obtenir le plus de récompenses possibles, il s'améliore progressivement. En guise d'exemple, on peut citer le programme AlphaGo qui a triomphé du champion du monde de jeu de Go. Ce programme a été entraîné par renforcement.

Quels sont les principaux algorithmes de Machine Learning ?

Il existe une large variété d'algorithmes de Machine Learning. Certains sont toutefois plus couramment utilisés que d'autres. Tout d'abord, différents algorithmes sont utilisés pour les données étiquetées.

Les algorithmes de régression, linéaire ou logistique, permettent de comprendre les relations entre les données. La régression linéaire est utilisée pour prédire la valeur d'une variable dépendante base sur la valeur d'une variable indépendante. Il s'agirait par exemple de prédire les ventes annuelles d'un commercial en fonction de son niveau d'études ou de son expérience.

La régression logistique est quant à elle utilisée quand les variables dépendantes sont binaires. Un autre type d'algorithme de régression appelé machine à vecteur de support est pertinent quand les variables dépendantes sont plus difficiles à classifier.

Un autre algorithme Machine Learning populaire est l'arbre de décision. Cet algorithme permet d'établir des recommandations basées sur un ensemble de règles de décisions en se basant sur des données classifiées.

Par exemple, il est possible de recommander sur quelle équipe de football parier en se basant sur des données telles que l'âge des joueurs ou le pourcentage de victoire de l'équipe.

Pour les données non étiquetées, on utilise souvent les algorithmes de « clustering ». Cette méthode consiste à identifier les groupes présentant des enregistrements similaires et à étiqueter ces enregistrements en fonction du groupe auquel ils appartiennent.

Auparavant, les groupes et leurs caractéristiques sont inconnus. Parmi les algorithmes de clustering, on compte les K-moyennes, le TwoStep ou encore le Kohonen.

Les algorithmes d'association permettent quant à eux de découvrir des schémas et des relations dans les données, et à identifier les relations « si/alors » appelées « règles d'association ». Ces règles sont similaires à celles utilisées dans le domaine du Data Mining ou forage de données.

Enfin, les réseaux de neurones sont des algorithmes se présentant sous la forme d'un réseau à plusieurs couches. La première couche permet l'ingestion des données, une ou plusieurs couches cachées tirent des conclusions à partir des données ingérées, et la dernière couche assigne une probabilité à chaque conclusion.

Un réseau de neurones « profond » est composé de multiples couches cachées permettant chacune de raffiner les résultats de la précédente. On l'utilise dans le domaine du Deep Learning.

Qu'est-ce que le Deep Learning ?

Le Deep Learning est une branche du Machine Learning, mais il s'agit aujourd'hui de la plus couramment utilisée. Il s'agit d'une invention de Geoffrey Hinton, datée de 1986.

Cette technique est appelée réseau de neurones profond. Cette profondeur correspond au large nombre de couches de nœuds de calcul qui constituent ces réseaux et travaillent en collaboration pour traiter les données et délivrer des prédictions.

Ces réseaux de neurones s'inspirent directement du fonctionnement du cerveau humain. Les nœuds de calcul sont comparables aux neurones, et le réseau en lui-même s'apparente au cerveau.

À quoi sert le Machine Learning ? Cas d'usage et applications

Le Machine Learning alimente de nombreux services modernes très populaires. On peut citer comme exemple les moteurs de recommandations utilisés par Netflix, YouTube, Amazon ou Spotify.

Il en va de même pour les moteurs de recherche web comme Google ou Baidu. Le fil d'actualité des réseaux sociaux tels que Facebook et Twitter reposent sur le Machine Learning, au même titre que les assistants vocaux tels que Siri et Alexa.

Toutes ces plateformes collectent des données sur les utilisateurs, afin de mieux les comprendre et d'améliorer leurs performances. Les algorithmes ont besoin de savoir ce que regarde le spectateur, sur quoi clique l'internaute, et à quelles publications il réagit sur les réseaux. De cette manière, ils sont ensuite en mesure de proposer de meilleures recommandations, réponses ou résultats de recherche.

Un autre exemple est celui des voitures autonomes. Le fonctionnement de ces véhicules révolutionnaires repose sur le Machine Learning. Pour l'heure, toutefois, les performances de l'IA restent limitées dans ce domaine. Si elle parvient à se garer ou à suivre une voie sur l'autoroute, le contrôle complet d'un véhicule en agglomération est une tâche plus complexe ayant provoqué plusieurs accidents tragiques.

Les systèmes de Machine Learning excellent aussi dans le domaine des jeux. L'IA a d'ores et déjà surpassé l'humain au jeu de Go, aux échecs, au jeu de dames ou au shogi. Elle arrive aussi à triompher des meilleurs joueurs de jeux vidéo comme Starcraft ou Dota 2.

On utilise aussi le Machine Learning pour la traduction linguistique automatique, et pour la conversion du discours oral à l'écran (speech-to-text). Un autre cas d'usage est l'analyse de sentiment sur les réseaux sociaux, reposant également sur le traitement naturel du langage (NLP).

Le Machine Learning est aussi utilisé pour l'analyse et la classification automatique des images de radiographies médicales. L'IA se révèle très performante dans ce domaine, parfois même plus que les experts humains pour détecter des anomalies ou des maladies. Toutefois, elle ne peut pas encore remplacer totalement les spécialistes compte tenu des enjeux.

Plusieurs entreprises ont tenté d'exploiter le Machine Learning pour passer en revue les CV des candidats de manière automatique. Toutefois, les biais des données d'entraînement mènent à une discrimination systématisée à l'égard des femmes ou des minorités.

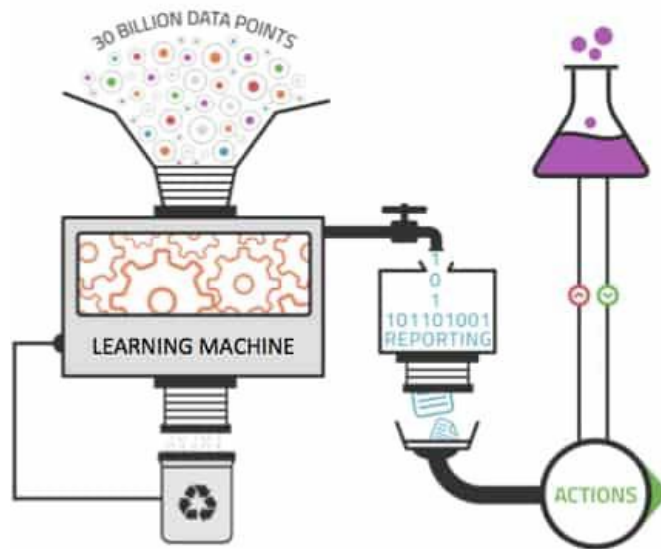
En effet, les systèmes de Machine Learning tendent à favoriser les candidats dont le profil est similaire aux candidats actuels. Ils tendent donc à perpétuer et à amplifier les discriminations déjà existantes dans le monde de l'entreprise.

C'est un réel problème, et Amazon a par exemple préféré cesser ses expériences dans ce domaine. De nombreuses entreprises tentent de lutter contre les biais dans les données d'entraînement de l'IA, telles que Microsoft, IBM ou Google.

La technologie controversée de reconnaissance faciale repose elle aussi sur le Machine Learning. Toutefois, là encore, les biais dans les données d'entraînement posent un grave problème.

Ces systèmes sont principalement entraînés sur des photos d'hommes blancs, et leur fiabilité se révèle donc bien inférieure pour les femmes et les personnes de couleur. Ceci peut mener à des erreurs aux conséquences terribles. Des innocents ont par exemple été confondus avec des criminels et arrêtés à tort...

Machine Learning et Big Data : pourquoi utiliser le Machine Learning avec le Big Data ?



Les outils analytiques traditionnels ne sont pas suffisamment performants pour exploiter pleinement la valeur du Big Data. Le volume de données est trop large pour des analyses complètes, et les corrélations et relations entre ces données sont trop importantes pour que les analystes puissent tester toutes les hypothèses afin de dégager une valeur de ces données.

Les méthodes analytiques basiques sont utilisées par les outils de business intelligence et de reporting pour le rapport des sommes, pour faire les comptes et pour effectuer des requêtes SQL. Les traitements analytiques en ligne sont une extension systématisée de ces outils analytiques basiques qui nécessitent l'intervention d'un humain pour spécifier ce qui doit être calculé.

Comment ça marche ?

Le Machine Learning est idéal pour exploiter les opportunités cachées du Big Data. Cette technologie permet d'extraire de la valeur en provenance de sources de données massives et variées sans avoir besoin de compter sur un humain. Elle est dirigée par les données, et convient à la complexité des immenses sources de données du Big Data. Contrairement aux outils analytiques traditionnels, il peut également être appliqué aux

ensembles de données croissants. Plus les données injectées à un système Machine Learning sont nombreuses, plus ce système peut apprendre et appliquer les résultats à des insights de qualité supérieure. Le Machine Learning permet ainsi de découvrir les patterns enfouis dans les données avec plus d'efficacité que l'intelligence humaine.

Des cours de Machine Learning sont disponibles sur le Web. Ils permettent notamment de débiter l'apprentissage automatique à partir du langage informatique Python. Ce dernier, assez simple à apprendre autorise donc les néophytes à tester des applications utilisant cette technique avec Python. De même, les open classroom permettent de découvrir gratuitement le fonctionnement de cette technique de traitement des données.

Machine Learning et Big Data : pourquoi le Machine Learning n'est rien sans Big Data

Sans le Big Data, le Machine Learning et l'intelligence artificielle ne seraient rien. Les données sont l'instrument qui permet à l'IA de comprendre et d'apprendre à la manière dont les humains pensent. C'est le Big Data qui permet d'accélérer la courbe d'apprentissage et permet l'automatisation des analyses de données. Plus un système Machine Learning reçoit de données, plus il apprend et plus il devient précis.

L'intelligence artificielle est désormais capable d'apprendre sans l'aide d'un humain. Par exemple, l'algorithme Google DeepMind a récemment appris seul à jouer à 49 jeux vidéo Atari. Par le passé, le développement était limité par le manque d'ensembles de données disponibles, et par son incapacité à analyser des quantités massives de données en quelques secondes.

Aujourd'hui, des données sont accessibles en temps réel à tout moment. Ceci permet à l'IA et au Machine Learning de passer à une approche dirigée par les données. La technologie est désormais suffisamment agile pour accéder aux ensembles de données colossaux et pour les analyser. De fait, des entreprises de toutes les industries se joignent désormais à Google et Amazon pour implémenter des solutions IA pour leurs entreprises.

Machine Learning et Big Data : les analyses prédictives donnent du sens au Big Data

Les analyses prédictives consistent à utiliser les données, les algorithmes statistiques et les techniques de Machine Learning pour prédire les probabilités de tendances et de résultats financiers des entreprises, en se basant sur le passé. Elles rassemblent plusieurs technologies et disciplines comme les analyses statistiques, le data mining, le modelling prédictif et le Machine Learning pour prédire le futur des entreprises. Par exemple, il est possible d'anticiper les conséquences d'une décision ou les réactions des consommateurs.

Les analyses prédictives permettent de produire des insights exploitables à partir de larges ensembles de données, pour permettre aux entreprises de décider quelle direction emprunter par la suite et offrir une meilleure expérience aux clients. Grâce à l'augmentation du nombre de données, de la puissance informatique, et du développement de logiciels IA et d'outils analytiques plus simples à utiliser, comme Salesforce Einstein, un grand nombre d'entreprises peuvent désormais utiliser les analyses prédictives.

Selon une étude menée par Bluewolf auprès de 1700 clients de Salesforce, 75% des entreprises qui augmentent leurs investissements dans les technologies analytiques en tirent profit. 81% de ces utilisateurs des produits Salesforce estime que l'utilisation des analyses prédictives est l'initiative la plus importante de leur stratégie de ventes. Les analyses prédictives permettent d'automatiser les prises de décision, et donc d'augmenter la rentabilité et la productivité d'une entreprise.

L'intelligence artificielle et le Machine Learning représentent le niveau supérieur des analyses de données. Les systèmes informatiques cognitifs apprennent constamment sur l'entreprise et prédisent intelligemment les tendances de l'industrie, les besoins des consommateurs et bien plus encore. Peu d'entreprises ont déjà atteint le niveau des applications cognitives, défini par quatre caractéristiques principales : la compréhension des données non structurées, la possibilité de raisonner et d'extraire des idées, la capacité à affiner l'expertise à chaque interaction, et la capacité à voir, parler et entendre pour interagir avec les humains de

façon naturelle. Pour cela, il convient de développer le traitement par algorithme des langages naturels.

Machine Learning et Big Data : l'apprentissage automatique au service du Data Management

Face à l'augmentation massive du volume de données stockées par les entreprises, ces dernières doivent faire face à de nouveaux défis. Parmi les principaux challenges liés au Big Data, on dénombre la compréhension du Dark Data, la rétention de données, l'intégration de données pour de meilleurs résultats analytiques, et l'accessibilité aux données. Le Machine Learning peut s'avérer très utile pour relever ces différents défis.

Toutes les entreprises accumulent au fil du temps de grandes quantités de données qui demeurent inutilisées. Il s'agit des Dark Data. Grâce au Machine Learning et aux différents algorithmes, il est possible de faire le tri parmi ces différents types de données stockées sur les serveurs. Par la suite, un humain qualifié peut passer en revue le schéma de classification suggéré par l'intelligence artificielle, y apporter les changements nécessaires, et le mettre en place.

Pour la rétention de données, cette pratique peut également s'avérer efficace. L'intelligence artificielle peut identifier les données qui ne sont pas utilisées et suggérer lesquelles peuvent être supprimées. Même si les algorithmes n'ont pas la même capacité de discernement que les êtres humains, le Machine Learning permet de faire un premier tri dans les données. Ainsi, les employés économisent un temps précieux avant de procéder à la suppression définitive des données obsolètes.

Cette technologie est aussi utile pour l'intégration de données. Pour tenter de déterminer le type de données qu'ils doivent agréger pour leurs requêtes, les analystes créent généralement un répertoire dans lequel ils placent différents types de données en provenance de sources variées pour créer un bassin de données analytique. Pour ce faire, il est nécessaire de développer des méthodes d'intégration pour accéder aux différentes sources de données en provenance desquelles ils extraient les données. Cette technique peut faciliter le processus en créant des

mappings entre les sources de données et le répertoire. Ceci permet de réduire le temps d'intégration et d'agrégation.

Enfin, l'apprentissage des données permet d'organiser le stockage de données pour un meilleur accès. Au cours des cinq dernières années, les vendeurs de solutions de stockage de données ont mis leurs efforts dans l'automatisation de la gestion de stockage. Grâce à la réduction de prix du SSD, ces avancées technologiques permettent aux départements informatiques d'utiliser des moteurs de stockage intelligents reposant sur le machine Learning pour voir quels types de données sont utilisés le plus souvent et lesquels ne sont pratiquement jamais utilisés. L'automatisation peut être utilisée pour stocker les données en fonction des algorithmes. Ainsi, l'optimisation n'a pas besoin d'être effectuée manuellement.

Mes sources :

Datascientest.com : <https://datascientest.com/machine-learning-tout-savoir>

ia-data-analytics.fr : <https://ia-data-analytics.fr/machine-learning/>

Wikipédia : https://fr.wikipedia.org/wiki/Apprentissage_automatique

Lebigdata.fr : <https://www.lebigdata.fr/machine-learning-et-big-data>

III- Le Machine Learning dans le futur

Certaines voix s'élèvent au sein des entreprises afin de rappeler que l'humanité est au début du développement de l'intelligence artificielle. Selon Alex Danvy, Évangéliste technique chez Microsoft France, le machine learning aujourd'hui est une forme simple d'IA. Les algorithmes ne sont pas encore capables d'accomplir les tâches aussi complexes que celles confiées à Skynet, le réseau informatique fictif du film Terminator. Qu'ils traitent des images, des sons, du texte, les algorithmes réalisent des tâches simples. Ce n'est qu'en interconnectant les algorithmes que l'on arrive à créer des systèmes plus intelligents. C'est de cette manière que sont pensées les voitures autonomes. Malheureusement, les acteurs de l'intelligence artificielle créent leurs solutions « dans leur coin », explique Alex Danvy. Selon lui, cela n'empêche pas l'émergence de solutions efficaces basés sur des algorithmes de machine learning « simples ».

Ce qu'explique Alex Danvy nous permet donc de penser que dans le futur si les acteurs de l'IA et donc du Machine Learning unissent leurs forces nous pourrions atteindre des intelligences dignes des plus grands films et donc de faciliter la tâche à beaucoup d'humain sur beaucoup de fonctionnalités différentes.

On croit être capable de grandes choses avec le Machine Learning d'aujourd'hui mais nous sommes seulement au commencement dans le futur ces intelligences seront capables de beaucoup plus.

IV- Conclusion

Pour conclure nous avons vu que plusieurs types de Machine Learning existe. Il y a l'apprentissage supervisé et l'apprentissage non-supervisé. Le premier consiste à s'entraîner avec des données déjà « étiquetées » alors que le second lui s'entraîne avec des données non « étiquetées ».

Il existe aussi différents types d'algorithmes tel que les algorithmes de régression, linéaire ou logistique, qui permettent de comprendre les relations entre les données. Un autre algorithme est l'arbre de décision, cet algorithme permet d'établir des recommandations basées sur un ensemble de règles de décisions en se basant sur des données classifiées. Les algorithmes de « clustering » sont utilisés pour les données non « étiquetées ». Les algorithmes d'association permettent quant à eux de découvrir des patrons et des relations dans les données. Il y a aussi les réseaux de neurones qui sont des algorithmes qui se présentent sous la forme d'un réseau à plusieurs couches.

Le Machine Learning grâce à ces différents algorithmes qui lui permettent de s'adapter à la situation est de plus en plus utilisé surtout par les grands services tel que Netflix, YouTube, Amazon ou bien Spotify ou bien de grands moteurs de recherche tel que Google. Il permet à ces grands services d'en apprendre plus sur l'utilisateur et donc de mieux le conseiller soit au niveau des services de chacun soit au niveau des pubs qui sont retransmises sur les services respectifs.

Cependant le Machine Learning commet encore des erreurs dans la reconnaissance faciale par exemple ou pour la conduite autonome en ville sur les voitures qui en sont équipées. Ces erreurs sont normales et seront corrigés dans le futur puisque nous sommes seulement au début de l'ère du Machine Learning et nous ne savons pas encore exploiter toutes les capacités de cet outil. Les Big Data sont donc là pour fortement aider au développement du Machine Learning et lui permettre de sans cesse s'améliorer.