République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Ammar Telidji – Laghouat –

Faculté de technologie

Département Mécanique

Mémoire pour l'Obtention du diplôme :

Master en Maintenance industrielle

Option :

Maintenance industrielle

Présenté par :

Benzita Samia

Thème :

Machine Learning Et Maintenance Prédictive

Soutenu devant le jury composé de :

Nom et Prénom	Grade	Qualité
Mr.Bensenouci Ahmida	MCA	Président
Mr.Rayane Karim	MCA	Encadreur
Mr.Bensahal Djamal	MCA	Examinateur

Année Universitaire : 2022/2023

Remerciements

Toute notre parfaite gratitude et remerciement à Allah le plus puissant qui m'a donné la force, le courage et la volonté pour élaborer ce travail.
C'est avec une profonde reconnaissance et considération particulière que je remercie mes professeurs Mr. Rayane Karim, et Mr. Bensenouci Ahmida, Mr. Bensahal Djamal
Pour leurs soutiens, leurs conseils judicieux et leurs grandes bienveillances durant l'élaboration de ce projet. Ainsi j'exprime ma reconnaissance à tous les membres de jury d'avoir accepté de lire ce manuscrit et d'apporter les critiques nécessaires à la mise en forme de cet ouvrage. Enfin, à tous ceux qui m'ont aidé de près ou de loin pour la réalisation de ce mémoire.

Dédicaces

A ma très chère mère, qui me donne toujours l'espoir de vivre et qui n'a jamais cessé de prier pour moi

A mon très cher père (يرحمو الله), pour ses encouragements, son soutien, surtout pour son amour

A mes sœurs Chahrazed, Nour,Siham

A mes frères Mohammed lamine, Aboubaker, Abdelrezzak

A mes meilleurs amis abdelkader, Azzedine,Nadjah, Wiam, Khaoula, Ouafaa

Et tout qui m'ont aidé et compulsé ce modeste travail.

Benzita Samia

Sommaire

Liste des Figures

Liste des Tableaux

Liste des abréviations

ML	Machine Learning
IA	Intelligence Artificielle
MI	Maintenance industrielle
MP	Maintenance prédictive
MS	Maintenance Systématique
MC	Maintenance Corrective
MCN	Maintenance Conditionnelle
SVM	Support Vector Machine
TTF	Le Temps De Fonctionnement Avant Panne
KNN	K-plus proche voisin

INTRODUCTION GENERALE

Introduction générale

Dans le domaine industriel, la maîtrise des coûts de production tout en assurant un niveau de qualité désiré constitue un défi clé dans ce domaine. L'un des objectifs majeurs des industriels est de minimiser au maximum les coûts élevés induits par les indisponibilités non planifiées et les pannes des équipements qui peuvent générer des pertes considérables. La maintenance prédictive est une méthode de maintenance proactive qui utilise des données et des techniques d'analyse pour prédire les défaillances potentielles des équipements avant qu'elles ne se produisent. Elle permet aux industriels de planifier les interventions de maintenance au moment opportun, avant que les équipements ne tombent en panne, ce qui réduit les temps d'arrêt imprévus et les coûts de maintenance élevés. La maintenance prédictive utilise une variété de technologies telles que la surveillance des vibrations, l'analyse thermique, l'analyse de l'huile, la surveillance de l'état des équipements, la surveillance de la consommation d'énergie, l'analyse des données de maintenance et d'autres méthodes pour surveiller l'état de santé des équipements. Ces données sont ensuite analysées pour prédire les défaillances potentielles et planifier les interventions de maintenance nécessaires. En utilisant la maintenance prédictive, les industriels peuvent également éviter les coûts inutiles liés au remplacement de pièces de rechange non défectueuses, ce qui réduit les coûts de maintenance globaux. En outre, cette méthode permet également de réduire les risques de dégâts matériels et humains imprévus, car elle permet d'identifier les problèmes avant qu'ils ne causent des dommages importants.En somme, la maintenance prédictive est un élément clé pour améliorer la fiabilité, la sécurité et la disponibilité des équipements de production tout en réduisant les coûts de maintenance. Elle permet également aux industriels de se concentrer sur des stratégies de production efficaces et durables, ce qui leur permet de rester compétitifs dans un marché mondial de plus en plus exigeant. La machine learning (apprentissage automatique) est une branche de l'intelligence artificielle qui permet aux systèmes informatiques d'apprendre et d'améliorer leur performance à partir de données, sans être explicitement programmés. Dans le domaine de la maintenance prédictive, le machine learning peut être utilisé pour améliorer la précision des prévisions de défaillance et pour identifier les modèles de défaillance complexes qui ne sont pas facilement détectables avec des méthodes traditionnelles.

Le machine learning est utilisé dans la maintenance prédictive pour développer des modèles de prédiction de défaillance en utilisant des données de capteurs et d'autres sources de données pertinentes. Ces modèles peuvent être formés à partir de données historiques pour prédire les défaillances futures avec une précision accrue. Par exemple, un modèle de machine learning peut être entraîné pour prédire les défaillances d'une machine en analysant les données de capteurs telles que les niveaux de vibrations, les températures, les niveaux d'huile, etc. Le modèle peut être entraîné pour reconnaître les modèles de défaillance complexes et pour alerter les équipes de maintenance avant que la défaillance ne se produise. Le machine learning peut également être utilisé pour optimiser les stratégies de maintenance en analysant les données de maintenance, telles que les temps de réparation et de remplacement de pièces, et en utilisant ces données pour améliorer la planification des interventions de maintenance. En utilisant le machine learning dans la maintenance prédictive, les industriels peuvent améliorer l'efficacité de leurs équipements et réduire les coûts de maintenance. En outre, cela peut également améliorer la sécurité des travailleurs en réduisant les risques de défaillance imprévus des équipements.

La maintenance traditionnelle, basée sur le diagnostic après-panne, a été remplacée par des approches plus prédictives telles que la maintenance préventive, la maintenance conditionnelle et la maintenance prédictive. La maintenance prédictive se concentre sur la surveillance en temps réel des équipements, la collecte de données et l'analyse de ces données pour prédire les défaillances avant qu'elles ne se produisent. L'apprentissage automatique, en tant que branche de l'intelligence artificielle, offre une méthode pour l'analyse des données et l'identification des modèles qui peuvent être utilisés pour améliorer la maintenance prédictive. Les algorithmes de machine learning peuvent être entraînés à partir de données historiques pour identifier des modèles de défaillance complexes et pour prédire avec précision les défaillances futures. En utilisant la maintenance prédictive et le machine learning, les entreprises peuvent optimiser leurs stratégies de maintenance, réduire les temps d'arrêt non planifiés et augmenter la durée de vie des équipements. En fin de compte, cela peut aider les entreprises à économiser de l'argent, à améliorer leur efficacité et à rester compétitives sur le marché.

Dans le premier chapitre est consacré à la maintenance industrielle, en mettant l'accent sur la maintenance prédictive. Nous abordons les différents enjeux de la maintenance dans l'industrie, qui ont évolué avec le temps et se reflètent dans les différentes formes de maintenance utilisées, depuis la maintenance traditionnelle jusqu'à la maintenance prédictive. Nous présentons également le lien entre la maintenance prédictive et l'apprentissage automatique, en explorant le concept de corrélation entre ces deux domaines.

Dans le chapitre 2 nous allons fournir une vue d'ensemble de l'intelligence artificielle en général, avec un accent particulier sur l'apprentissage automatique. Nous aborderons les techniques les plus couramment utilisées dans la littérature récente, en définissant l'apprentissage automatique et en présentant ses types les plus courants, tels que la régression et la classification. Nous mettrons également l'accent sur les algorithmes liés à la prédiction des défaillances et à la maintenance. L'objectif de ce chapitre est de présenter une approche d'apprentissage automatique qui sera utilisée pour la modélisation et l'implémentation dans les chapitres suivants.

Dans le chapitre 3 nous présentons une étude de cas en utilisant et nous apporterons notre contribution dans le but de résoudre le problème en question. Après la collecte et la description de l'ensemble de données utilisé (Des filtres à air pour véhicules roulants), nous procédons au prétraitement. La sélection d'une approche d'apprentissage automatique appropriée est cruciale sur le plan analytique. Ensuite, nous présentons un certain nombre d'algorithmes adoptés pour modéliser l'objectif initial, ainsi que les techniques permettant de mesurer les performances des modèles construits. En expliquant la mise en œuvre de l'approche, en commençant par une vue d'ensemble des outils et de l'environnement de développement. Ensuite, nous présenterons quelques interfaces illustrant les résultats obtenus en appliquant les algorithmes choisis sur l'ensemble de données préparés. Nous validons enfin cette approche avec des métriques de performance. Dans la conclusion générale, nous récapitulons l'ensemble des développements par rapport à l'objectif initial de l'étude, nous résumons les principaux résultats obtenus.

Chapitre 1 : Maintenance Prédictive

1.1 Introduction

La maintenance est une tâche très importante dans un monde industriel pour le bon fonctionnement des différents équipements.la maintenance des outils s'impose donc avec évidence comme une nécessité et ce à toutes les étapes de la durée de vie d'un système, d'un mécanisme ou d'un organe outil. Dans ce chapitre nous parlerons tout d'abord de ce qu'est la maintenance industrielle ainsi que les types de la maintenance industriel et la relation entre la maintenance prédictive et Machine Learning et les Avantages et inconvénients de la maintenance prédictive, les outils et sans oublier de donner quelques exemples de domaines auxquelles elle s'applique.

1.2 La maintenance industrielle

Une première définition normative de la maintenance fut donnée par l'AFNOR :« l'ensemble des actions permettant de maintenir ou de rétablir un bien dans un état spécifié ou en mesure d'assurer un service déterminé » [1]. La maintenance industrielle définie selon la norme AFNOR :« est l'ensemble de toutes les actions techniques, administratives et de management durant le cycle de vie d'un bien, destinées à le maintenir ou à le rétablir dans un état dans lequel il peut accomplir la fonction requise » [2]. Elle comprend ainsi ensemble d'actions de dépannage, de réparation, de contrôle et inspection et révision et diagnostic et vérification des équipements matériels etc…, pour le principale objectif de zéro panne et d'assurer le rendement global maximum. Enfin par l'amélioration du matériel (pièces de rechange et outillage). Une bonne maintenance et une bonne prévention, par des réparations rapides et efficaces donnés optimisant les coûts (plan économique) et techniquement la disponibilité, la fiabilité, la maintenabilité.

1.3 Les types de maintenance

Il existe 02 types de maintenance industrielle principaux sont : **la maintenance corrective** est effectuée après défaillance et tout équipement (la maintenance curative {réparation}. Et la maintenance palliative {dépannage}), **la maintenance préventive** est effectuée dans l'intention de réduire la probabilité de défaillance d 'un bien ou d 'un service rendu (la maintenance préventive systématique, la maintenance préventive conditionnelle, la maintenance prévisionnelle (prédictive). Nous allons dans cette partie du mémoire présenter les principaux types de maintenance industrielle.

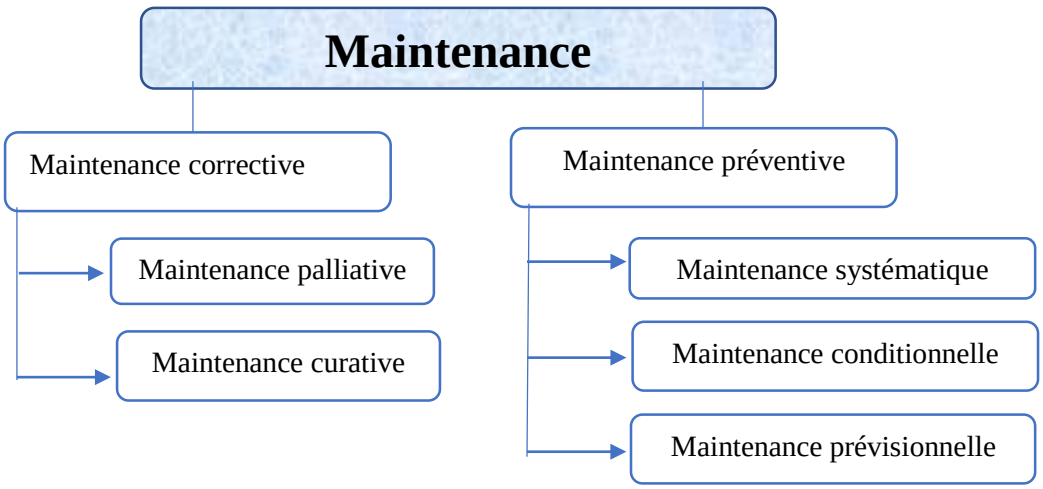


Figure 1-1 les différents types de maintenance .[03]

1.3.1 Maintenance corrective

La maintenance corrective peut être définie comme « Opération de maintenance effectuée après défaillance » [04] exécutée après la détection d'une panne, qui intervient après défaillance, maintenance subie. C'est l'ensemble des activités réalisées après défaillance d'un bien ou dégradation de sa fonction, afin de lui permettre d'accomplir, au moins provisoirement, une fonction requise, peut être divisée en deux typés de maintenance : maintenance palliative (dépannage), maintenance curative (réparation) d'amélioration corrective.

1.3.2 Maintenance préventive

Selon la norme la définit comme suit : « Maintenance exécutée à des intervalles prédéterminés ou selon des critères prescrits et destinée à réduire la probabilité de défaillance ou la dégradation du fonctionnement d'un bien ». [05] Maintenance exécutée dans l'intention de réduire la probabilité de défaillance ou la dégradation d'un bien ou d'un service rendu. L'objectif de la maintenance préventive :

- Demeure de réduire la probabilité de défaillances en service.
- Augmenter la durée de vie des matériels.
- Diminuer le temps d'arrêt en cas de révision ou de panne.
- Diminuer le coût de la maintenance.
- Supprimer les causes d'accidents graves.[06]

1.3.3 Maintenance préventive systématique

La maintenance systématique peut être définit comme suit : est effectuée selon un échéancier, en fonction du temps ou du nombre d'unités d'usage, mais sans contrôle préalable du bien. [07] Maintenance programmée, organisé, planifiée, par potentiel d'heures. Est exécutée à intervalles

fixes et prédéfinis, on parle de la maintenance systématique (MS). Ce type de maintenance est déclenché selon un calendrier (heures de travail, kilomètres effectués, etc.) et se traduit par le remplacement périodique de pièces, sans contrôle préalable et quel que soit l'état de dégradation des équipements. La maintenance préventive systématique peut conduire à du sur-entretien c'est à dire à un excès d'interventions inutiles, et donc à des gaspillages financiers pour l'entreprise
La maintenance systématique peut être appliquée pour :

- Équipements ayant un coût de défaillance élevé.
- Équipements soumis à une législation impérative (le domaine du nucléaire, l'aéronautique, les appareils sous pression, les chaudières…etc.).
- Équipements dont une défaillance met en cause la sécurité du personnel.
- Composants et sous-ensembles dont les durées de vie sont bien connus.

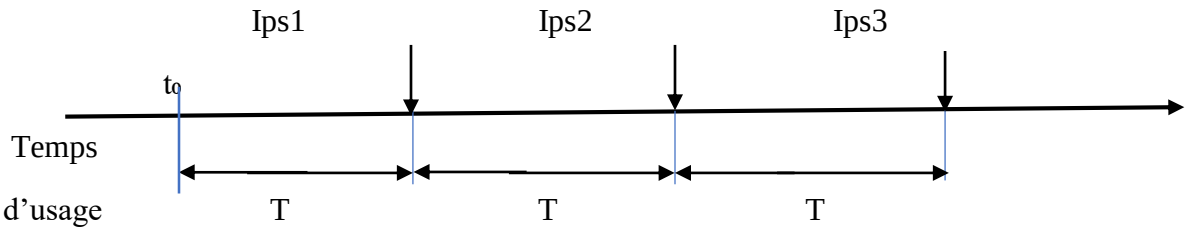


Figure 1-2 intervention de la maintenance préventive systématique.[08]

La figure 1-2 : Illustre l'intervention préventive systématique (Ipsi) effectuée après un échéancier d'intervention prédéterminée (T) constante.

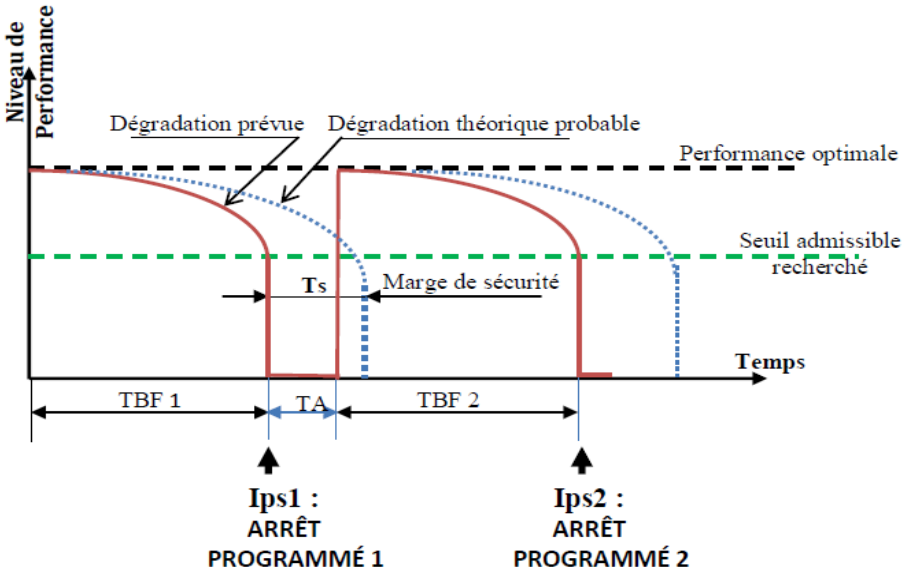


Figure 1-3 principe de la maintenance préventive systématique.[08]

La figure 1-3 illustre le graphe caractéristique de la maintenance préventive systématique et montre que l'intervention est programmée selon un échéancier (période) prédéfinie.

L'écart entre la dégradation théorique probable et celle prévue par l'exploitant représente « la marge de sécurité » que l'en note par {TS}.

- Plus la marge de sécurité (Ts) est importante plus l'exploitant risque une panne imprévue.
- Plus la marge de sécurité (Ts) est faible plus l'exploitant échange une pièce presque neuve, ce qui entraînera un gaspillage de potentiel (pièce) plus important. Nous verrons au paragraphe suivant « la maintenance conditionnelle » un meilleur moyen de résoudre ce problème.

La périodicité (TBFi) d'intervention systématique est déterminée par le service maintenance en exploitant l'historique des défaillances.

A un certain seuil admissible recherché (correspondant à un échéancier TBFi constant), un arrêt programmé par le service maintenance pour intervention systématique programmée.

TA : Temps d'arrêt pour une intervention de maintenance systématique [08].

1.3.4 Maintenance préventive conditionnelle

La maintenance conditionnelle, on condition, selon l'état ou prédictive. Selon la norme AFNOR X 60-000 « Les activités de maintenance conditionnelle sont déclenchées suivant des critères prédéterminés significatifs de l'état de dégradation du bien ou du service. » [09] Maintenance subordonnée à un type d'événement prédéterminé (mesure, diagnostic).

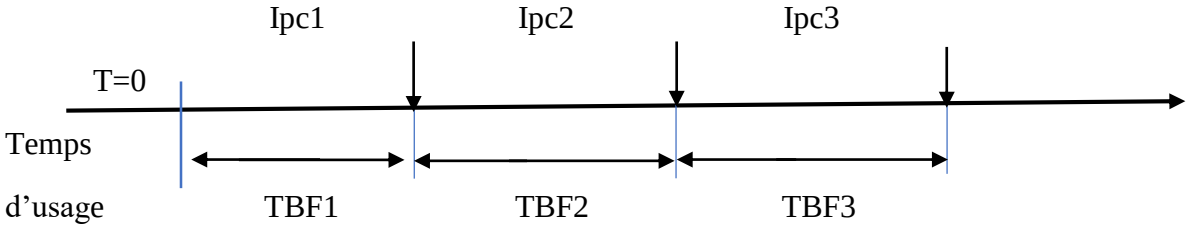


Figure 1-4 Intervention de la maintenance préventive conditionnelle. [08]

Avec :

Ipci : Intervention de maintenance préventive conditionnelle.

TBFi : Temps de Bon Fonctionnement avant l'arrêt programmé pour Ipci.

TBF1≠ TBF2≠TBF3 à cause des conditions d'exploitation qui ne sont pas identiques.

La figure 1-4 illustre le principe de la maintenance préventive conditionnelle. Elle se rapporte au suivi par mesures périodiques (surveillance) ou en temps réel à travers des capteurs d'une dégradation jusqu'au seuil d'alarme qui déclenche une intervention préventive conditionnelle.

L'intervention Ipci sera programmée à partir de l'alarme, suivant un temps de « réaction » du service maintenance à prédéterminer.

Le temps de réaction est le temps nécessaire pour préparer et programmer l'intervention à savoir : Qui (nom du technicien), Quand (date et heure d'intervention), avec quels moyens (outillages et pièces de rechanges….etc nécessaires pour l'intervention) et ce sans dépasser le seuil admissible.

TA c'est le temps d'arrêt machine.

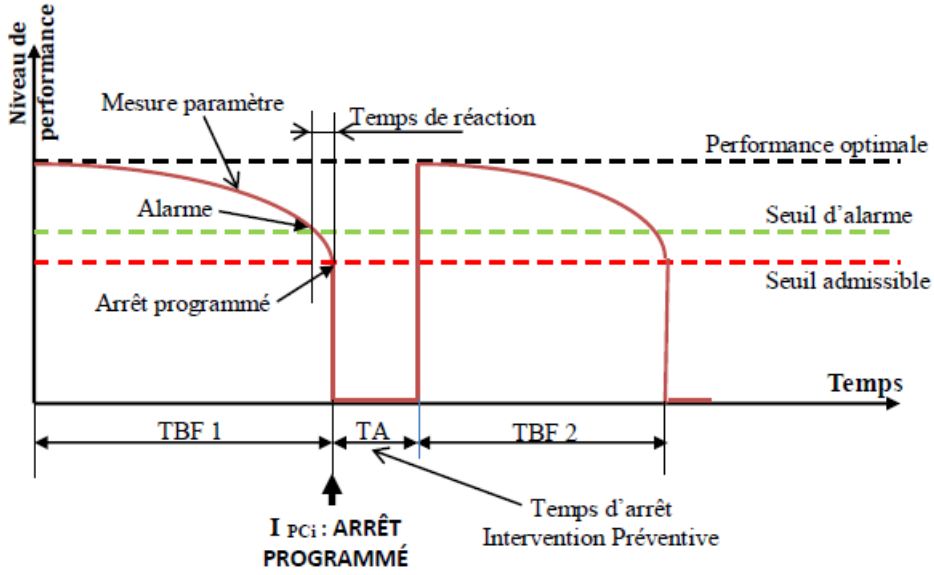


Figure 1-5 principe de la maintenance préventive conditionnelle.[08]

- Seuil admissible : Il sera choisi en fonction de contraintes réglementaires lorsqu'elles existent. Exemple : le changement conditionnel du pneu usé de votre voiture ! Sinon, il est rare que les constructeurs fournissent des préconisations. La démarche expérimentale est donc souvent utilisée.

- Seuil d'alarme : Il sera choisi à partir du seuil d'admissibilité, en prenant en compte :

- La vitesse de dégradation.

- Le temps de réaction avant l'intervention [08].

1.3.5 La maintenance prédictive

La maintenance prédictive ou "Maintenance Prévisionnelle ". Elle est définie selon la norme comme : « maintenance conditionnelle exécutée suite à une prévision obtenue grâce à une analyse répétée ou à des caractéristiques connues et à une évaluation des paramètres significatifs de la dégradation du bien » [09]. La maintenance prédictive est tout à fait appropriée avec la maintenance préventive dans la mesure où il s'agit d'intervenir sur les machines avant même qu'elles ne tombent en panne. Toutefois, les deux approches sont à distinguer. La maintenance préventive consiste à prévenir les risques de panne des machines et des engins en programmant des interventions de maintenance selon un calendrier établi en fonction de paramètres figés [10]. La maintenance préventive permet de réduire la fréquence à laquelle la maintenance d'une machine est effectuée et d'éviter une maintenance réactive imprévue sans entraîner de coûts élevés liés à une maintenance trop importante Il est nécessaire que toute stratégie de maintenance

minimise les taux de défaillance de l'équipement, améliore l'état de l'équipement, prolonge la durée de vie de l'équipement et réduise les coûts de maintenance. [11]

1.4 Maintenance prédictive et Machine Learning

La maintenance prédictive ou prévisionnelle est le processus qui consiste à utiliser l'analyse des données, l'apprentissage automatique, l'intelligence artificielle et d'autres techniques de gestion des données pour prédire quand des équipements industriels sont susceptibles de tomber en panne. L'objectif de la maintenance prédictive est de détecter la défaillance imminente des actifs avant qu'ils ne tombent en panne, notamment grâce à des systèmes de seuils et d'alertes. Autrement dit, elle s'attaque directement à l'une des pertes de productivité pure pour un industriel, à savoir ses temps d'arrêt machine. [12] Le Machine Learning est une branche de l'intelligence artificielle qui permet, grâce à des algorithmes d'apprentissage automatique, d'analyser des données et de diagnostiquer des pannes à un stade précoce. [13]

1.5 Les technologies de l'industrie au cœur de la maintenance prédictive

Les capteurs IIoT pour capter les données : La première étape de la maintenance prédictive consiste à collecter des données sur l'équipement. Ces données peuvent être recueillies par des capteurs installés sur l'équipement, tels que des capteurs de pression sur une chaudière, des capteurs de vibration sur un moteur ou des capteurs de tension sur une installation électrique. Les capteurs IIoT (Industrial Internet of Things ou Internet Industriel des Objets) étant reliés au réseau, toutes les données mesurées sont envoyées vers un serveur ou dans le cloud afin d'être exploitées par une solution logicielle qui effectuera l'analyse des données, indispensable pour la maintenance prédictive. [12]

Figure 1-6 Capteurs [12]

1.6 Application de la maintenance prédictive dans l'industrie

La maintenance prédictive domine particulièrement le marché de l'industrie car elle s'applique notamment aux machines tournantes. La crise sanitaire du coronavirus a par ailleurs accéléré le déploiement de la maintenance prédictive dans l'industrie, afin d'éviter aux techniciens de se déplacer inutilement.

1.6.1 La maintenance prédictive des automobiles

Avec l'essor des véhicules connectés, qui permettent de collecter de nombreuses informations, la maintenance prédictive suscite l'intérêt du secteur automobile, désireux de connaître en temps réel l'usure des pièces mécaniques des parcs automobiles. Par exemple, avec son offre Michelin Tire Care, le fabricant français assure la maintenance prédictive de ses pneumatiques sous forme de service à ses clients.

1.6.2 La maintenance prédictive des ascenseurs

Vous faites partie de celles et ceux qui ont la hantise de se retrouver bloqué dans un ascenseur ? Avec la maintenance prédictive et l'IoT, il devient possible de réparer une panne d'ascenseur avant que celle-ci ne se produise. Les principaux acteurs du marché de l'ascenseur se sont ainsi lancés dans une course aux services de maintenance prédictive pour leurs équipements, avec des ascenseurs toujours plus connectés et reliés à l'intelligence artificielle. A la clé, une résolution des problèmes avant qu'ils ne surviennent, et une sécurité optimale pour les usagers. La start-up parisienne WeMaintain a observé une division par trois des pannes au bout de 6 mois de maintenance prédictive. Plus de 10% du parc géré par l'ascensoriste finlandais Kone était connecté en juin 2020. Le groupe prévoit de connecter l'ensemble de ses nouveaux appareils pour proposer des services IoT d'ici 2022. [13]

1.7 Avantages et les inconvénients de la maintenance prédictive

1.7.1 Avantages

- Limiter la rupture de la chaîne de production.
- D'allonger la durée de vie des équipements.
- D'améliorée la fiabilité des équipements permettant ainsi d'optimiser la production.
- L'amélioration du taux de rendement global.
- De diminuer le nombre d'interruptions des machines pour des opérations de maintenance.
- De diminuer le nombre de pannes.
- De mieux planifier les interventions.
- De mieux préparer les équipes d'intervention.
- De mieux échanger entre les professionnels de maintenance et les équipes de production.
- De mieux anticiper et gérer les besoins de pièces détachées des outils. [14]

1.7.2 Inconvénients

- Les données peuvent être mal interprétées, entraînant de fausses demandes de maintenance.
- Il est coûteux de mettre en place un système IoT complet.
- L'analyse prédictive peut ne pas prendre en compte les informations contextuelles, Comme l'âge de l'équipement ou les conditions météorologiques.
- La maintenance prédictive peut décourager l'inspection physique proactive Et entretien des équipements.
- Les activités de maintenance préventive peuvent être déclenchées par des échéanciers plutôt Que l'état de la machine d'origine. [15]

1.8 Conclusion

Dans ce chapitre nous avons présenté dans la première partie une définition et une vue générale sur la maintenance industrielle. Puis nous avons mentionné les types de maintenance et la relation entre la maintenance prédictive et Machine Learning et les technologies de l'industrie au cœur de la maintenance prédictive et domaine application de la maintenance prédictive dans l'industrie aussi les avantages et les inconvénients de la maintenance prédictive.

Chapitre 2 : Machine Learning

2.1 Introduction

L'apprentissage automatique, également connu sous le nom de « machine learning » est une branche de l'intelligence artificielle (IA) et de l'informatique qui consiste d'apprendre à partir de données et des algorithmes pour imiter la façon dont les êtres humains apprennent, en améliorant progressivement leur précision. L'apprentissage automatique est une composante clé de la science des données et joue un rôle important dans de nombreux domaines. A partir l'utilisation de méthodes statistiques et des algorithmes sont entraînés à effectuer des classifications ou des prévisions, ce qui permet de découvrir des informations précieuses dans le cadre de projets d'exploration des données. Ces informations peuvent être utilisées pour prendre des décisions dans les applications et les entreprises, et ont idéalement un impact sur les principales métriques de croissance. Avec la croissance continue des données et la numérisation de nombreuses industries, la demande pour des spécialistes des données compétents ne fera qu'augmenter dans les années à venir. Ces professionnels joueront un rôle essentiel dans l'identification des questions commerciales les plus pertinentes et la collecte des données nécessaires pour y répondre. [16] Dans ce chapitre nous exposons tout d'abord une brève histoire de l'intelligence artificielle et du machine learning. Ensuite, nous présentons les concepts et les techniques les plus importantes utilisées en machine learning. Enfin, nous établirons un état de l'art sur l'application du machine learning à la maintenance prédictive.

2.2 Définition de l'intelligence artificielle

En d'autres termes, l'intelligence artificielle est une technologie qui permet aux machines d'effectuer des tâches qui nécessitent normalement l'intelligence humaine, telles que la reconnaissance de voix, la prise de décisions et la résolution de problèmes complexes. Les systèmes d'IA peuvent apprendre à partir de données, à travers des algorithmes d'apprentissage automatique, et s'améliorer progressivement en fonction de l'expérience acquise. L'intelligence artificielle peut prendre différentes formes, telles que les chatbots, les systèmes de recommandation, la reconnaissance d'images et la détection de fraudes. Elle peut être utilisée dans de nombreux secteurs d'activité, tels que la finance, la santé, l'industrie, l'éducation et les transports, pour améliorer l'efficacité et la précision des processus. Bien que certains craignent que l'IA ne remplace les emplois humains, elle est en réalité conçue pour travailler en collaboration avec les êtres humains, en leur permettant de se concentrer sur des tâches à plus forte valeur ajoutée. L'intelligence artificielle est ainsi un atout précieux pour les entreprises qui cherchent à améliorer leur performance et leur compétitivité. [17]

2.3 Définition du machine learning

Le Machine Learning est une sous-branche de l'Intelligence Artificielle qui permet aux machines d'apprendre à partir de données sans être explicitement programmées pour chaque tâche spécifique. L'IA a pour objectif de reproduire certains aspects du comportement humain tels que la capacité à comprendre un langage naturel, à raisonner, à percevoir, à apprendre et à prendre des décisions. L'IA peut être construite de différentes manières, notamment en utilisant des approches symboliques ou neuronales. Actuellement, l'IA neuronale est la méthode la plus

couramment utilisée en raison de sa capacité à s'adapter et à apprendre à partir des données. Le Machine Learning nécessite souvent de grandes quantités de données pour s'entraîner et s'améliorer, ce qui soulève des questions éthiques liées à la protection de la vie privée et à la sécurité des données.[18] La Prédictive Maintenance Usine Machine Learning est une application de l'IA qui permet de détecter les défaillances potentielles des équipements industriels en analysant les données de performance et en prédisant les pannes avant qu'elles ne se produisent. Cela permet aux entreprises de réduire les temps d'arrêt imprévus, d'améliorer la sécurité des travailleurs et de réduire les coûts de maintenance.[19] Durant ces dernières années, le machine learning est devenu de plus en plus important dans le domaine informatique car les données peuvent être collectées et stockées beaucoup plus facilement.

2.4 Historique

L'histoire de l'apprentissage machine est souvent associée à celle de l'intelligence artificielle. L'apprentissage machine, considéré comme un moyen technique qui s'inspire de l'intelligence biologique, a poussé, dès ses débuts, les scientifiques à toujours miser sur la création d'ordinateurs dotés d'une capacité d'apprentissage et d'autoadaptation.

1950

Les fondements de l'intelligence artificielle remontent aux années 1950, alors qu'Alan Turing, fondateur du test de Turing, mesure la capacité d'un ordinateur à interagir comme un être humain. Selon Turing, les ordinateurs sont considérés comme étant intelligents dès qu'ils peuvent tenir une conversation rationnelle avec un opérateur humain sans que celui-ci ne se rende compte qu'il interagit avec un ordinateur. À l'époque de Turing, les technologies informatiques étaient à un stade embryonnaire, les ordinateurs possédaient des capacités de mémoire et de calcul extrêmement faibles.

La création du perceptron (1958)

En 1958, Frank Rosenblatt du laboratoire aéronautique de l'université Cornell, présente le Perceptron, le premier algorithme d'apprentissage supervisé basé sur les réseaux de neurones artificiels. Cet algorithme est capable de reconnaître des formes élémentaires. Il procède à une simple classification binaire, par exemple classer une nouvelle forme en une forme X ou Y.

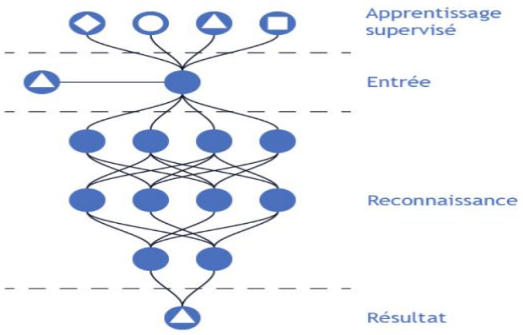


Figure 2-1 La création du perceptron (réseaux de neurones artificiels).[20]

1959 Le terme apprentissage machine popularisé

Le terme apprentissage machine devient populaire en 1959 grâce aux travaux du chercheur Arthur Samuel. Il présente un jeu de dames capable d'apprendre à mesure que le jeu évolue à partir de données historiques stockées dans sa mémoire.

1986 L'invention des algorithmes de rétro propagation

La publication de David Everett Rumelhart, Geoffrey Everest Hinton et Ronald Williams intitulé « Learning Representations by Back-Propagating Errors », dans la revue scientifique Nature, est considérée comme une découverte majeure dans la discipline de l'apprentissage machine. Il s'agit d'une méthode de rétro propagation d'un signal permettant à un réseau de neurones artificiels de corriger ses prédictions et d'ajuster automatiquement ses fonctions. Avec cette avancée, le progrès de la discipline atteint un certain plateau, les percées scientifiques se raréfient et, de ce fait, l'intérêt de la communauté scientifique à cet égard se dissipe.

2012 Le développement d'un algorithme de reconnaissance d'images performant

En 2012, le professeur Geoffrey Hinton et deux de ses étudiants, Alex Krizhevsky et Ilya Sutskever, réussissent à entraîner un algorithme de reconnaissance d'images capable d'identifier automatiquement les objets d'une image à un taux d'exactitude jamais atteint auparavant, soit 84,7 % à l'aide de processeurs graphiques.

2013 à aujourd'hui La multiplication des librairies d'information ouverte

Plusieurs grandes entreprises (Google, Microsoft, IBM, etc.) et d'importants laboratoires universitaires (Université de Montréal, University of Toronto, Stanford, MIT, etc.) travaillent d'arrache-pied pour démocratiser la discipline de l'apprentissage machine en mettant à la disposition des chercheurs et des développeurs des librairies d'informations ouvertes (open source) telles que TensorFlow, PyTorch, etc. [20]

2.5 Cas d'applications du machine Learning

Avec la démocratisation des algorithmes d'apprentissage machine et la disponibilité des données nécessaires à leur bon fonctionnement, force est de constater l'utilisation croissante de cette technologie dans différents domaines, Le machine learning concerne tous les secteurs d'activité, notamment l'industrie, Juridique, transport-luxe, comptabilité, le commerce, la santé…

Dans le monde de l'industriel :

- Maintenance prédictive et surveillance des équipements.
- L'analyse de données permet également d'optimiser les cycles de production complexes de certaines usines.
- Prévision des stocks.
- Analyse de marché/compétition.

Marketing :

- Amélioration des compagnes de pubs en faisant des analyses des précédentes pubs qui ont eu du succès.
- Augmenter les ventes en visant personnellement les clients qui ont le plus de chance de faire l'achat.

Les ressources sont humaines :

- Automatisation de la recherche de profils.
- CV correspondants.
- Construction de référentiels de concours.

Réseaux sociaux :

- Etudes de tendances.
- Etudes d'image de marques ou de produits.
- Analyse des réactions à l'actualité (opinions).

Les placements de produits comme le choix des publications dans votre fil d'actualité Facebook. En fonction de quelques critères comme les goûts de l'utilisateur, l'algorithme va déterminer les chances que vous aimiez la publication et décider ensuite s'il s'affiche ou non dans votre fil d'actualité.

E-commerce :

- Recherche des acheteurs basés sur le sens, gestion des synonymes.
- Analyse de sentiment dans les commentaires de site web ou forums et Evaluation de la satisfaction client.
- Utilisation pour la personnalisation de recommandation produit.[21]

2.6 Les approches de l'apprentissage automatique

Il existe 3 grandes approches en apprentissage automatique : l'apprentissage supervisé et l'apprentissage non supervisé et par renforcement chacun procédant par ces propres méthodes comme sur la figure 2-2. L'apprentissage supervisé est ce qui nous intéresse pour la prédiction. Elle est constituée de plusieurs algorithmes que ce soit pour la régression ou la classification. Pour choisir un algorithme il faut comprendre les fondements des algorithmes existants et de ce qui permet de les distinguer ce qui permet de créer les modèles qui traiteraient au mieux un problème particulier.

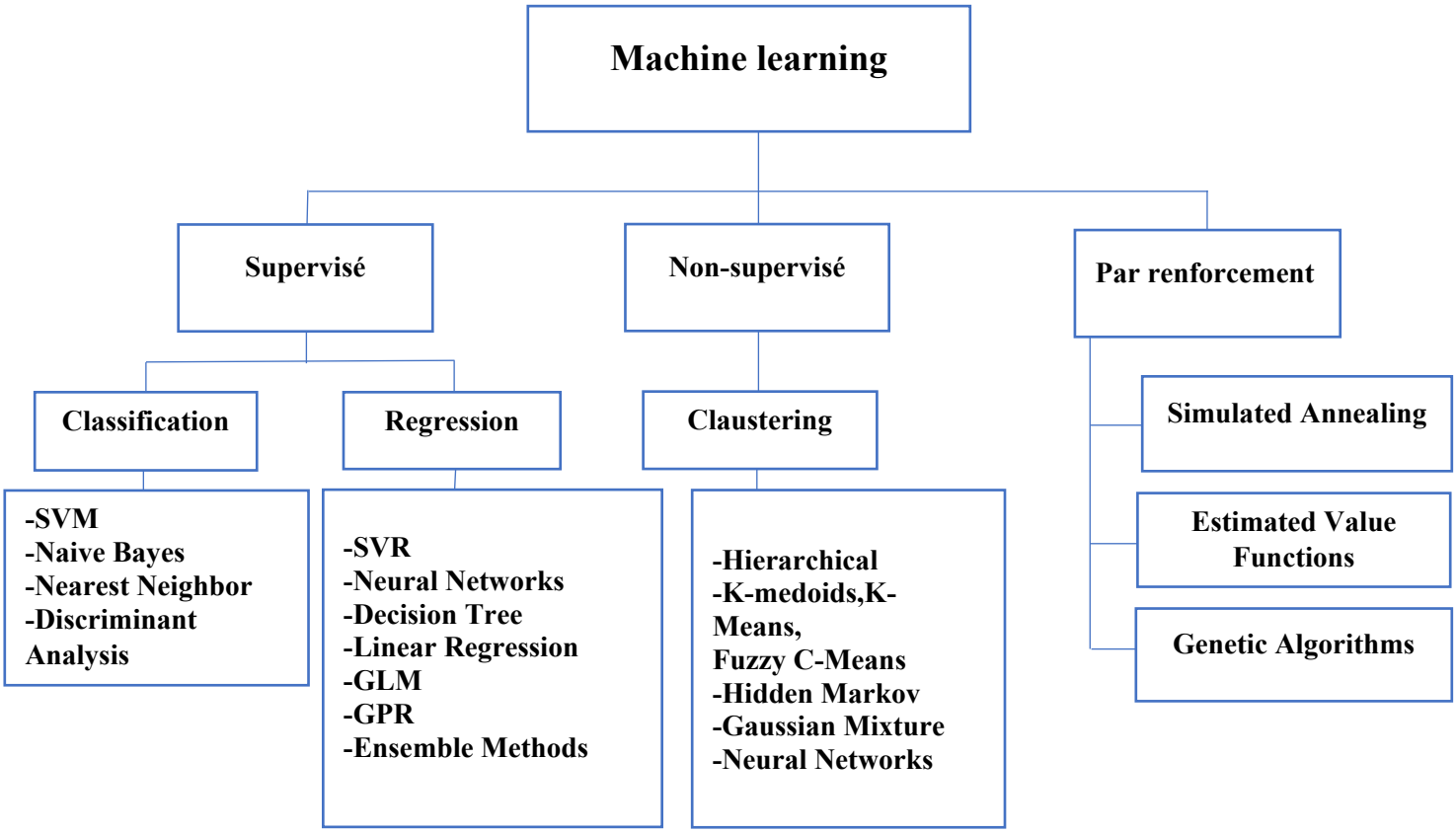


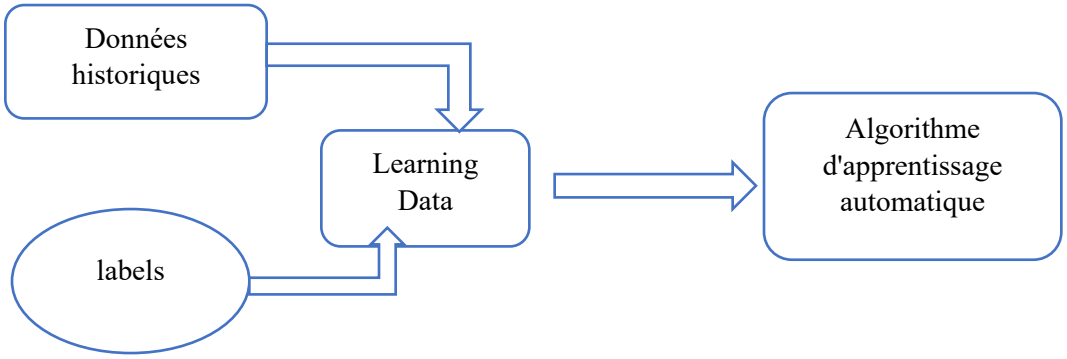
Figure 2-2 Classifications dans les techniques d'apprentissage automatique. [22]

Effectivement, il existe différentes façons pour une machine d'apprendre, et les approches de l'apprentissage automatique se distinguent généralement par les types de problèmes qu'elles tentent de résoudre et le type de données utilisé pour l'entraînement du modèle. Voici les trois sous-domaines principaux de l'apprentissage machine :

2.6.1 L'apprentissage supervisé

Cette méthode s'appuie sur un algorithme d'apprentissage supervisé, autrement dit que la machine s'intéresse à des données d'entrée et de réponses aux données déjà connues. Ce traitement des données connues va permettre à l'algorithme de se former et d'élaborer un modèle de prévisions adapté pour les données à traiter par la suite. Pour développer ces modèles prédictifs, l'apprentissage supervisé va notamment se baser sur des techniques de classification. La mise en pratique des techniques de classification implique le fait que les données peuvent être identifiées et catégorisées en fonction de leurs caractéristiques. L'application des techniques de classification se retrouvent dans plusieurs domaines : digital, médical, banque. Un des exemples de classification les plus courants et la détermination automatique de la nature d'un mail : authentique ou spam. L'apprentissage supervisé se retrouve également dans les techniques de régression. [23] Cette approche consiste à entraîner un modèle à partir d'exemples d'entrée et de

sortie (étiquettes) connus. Le modèle apprend ainsi à prédire les sorties correctes pour les nouvelles entrées qui lui sont présentées. Les exemples courants incluent la classification, la régression et la détection d'anomalies.[24] L'apprentissage supervisé est une méthode très riche, mais elle nécessite des données de qualité et une bonne compréhension du problème à résoudre. Il est important de noter que l'algorithme ne sera pas capable de généraliser au-delà des données d'entraînement, ce qui peut conduire à des prédictions erronées si les données d'entraînement ne sont pas représentatives des données réelles. Il est donc important de sélectionner les données avec soin et de s'assurer que le modèle est régulièrement mis à jour pour prendre en compte les changements dans les données réelles.



Apprentissage automatique supervisé Étape 1. [15]

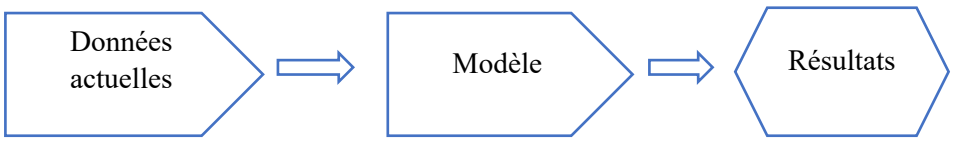


Figure 2-3 Apprentissage automatique supervisé Étape 2. [15]

L'apprentissage automatique supervisé peut se diviser en deux classes :

- **Classification :** la variable de sortie est une catégorie.
- **Régression :** la variable de sortie est une valeur spécifique. [25]

Quelles sont les applications de l'apprentissage supervisé :

Il existe de nombreuses applications de l'apprentissage supervisé :

- Le traitement automatique du langage.
- La reconnaissance vocale.
- La vision par ordinateur.
- La bio-informatique.

On utilise également l'apprentissage supervisé pour la détection de spams dans les mails, la gestion des chatbots et voicebots, ainsi que pour la robotique. Cette méthode permet aussi le développement des technologies embarquées dédiées aux véhicules autonomes. [26]

Quels sont les algorithmes d'apprentissage supervisé :

Pour créer un modèle d'apprentissage supervisé, on peut recourir à différents algorithmes :

- La régression linéaire : y = c + b * x.
- La régression logistique : h(x) = 1 / (1 + e^-x).
- L'arbre de décision avec différentes variables de sortie.
- la machine à vecteur de support (SVM).

On peut également évoquer l'algorithme k-NN pour réaliser des tests et la classification naïve bayésienne pour catégoriser des volumes de données importants. [26]

2.6.2 L'apprentissage non supervisé

A l'inverse de l'apprentissage supervisé, celui non supervisé va s'appuyer sur des données d'entrée dont les réponses ne sont pas identifiées. Ce type de l'objectif de technique est de mettre en exergue des modèles intrinsèques aux données traitées. Parmi l'apprentissage non supervisé, on retrouve la technique de clustering. Ce modèle d'apprentissage non supervisé est le plus courant et permet d'identifier des points communs entre certaines données et de les regrouper clairement par groupe. [23] Cette approche vise à découvrir des structures ou des modèles cachés dans les données sans étiquettes d'entraînement connues. Les exemples courants incluent la segmentation de données, la détection d'anomalies, la réduction de dimensionnalité et la clustering. [24] L'apprentissage non supervisé est une méthode très utile pour la découverte de connaissances à partir de données brutes, mais il peut être plus difficile à mettre en œuvre que l'apprentissage supervisé car il nécessite souvent un prétraitement des données pour les rendre exploitables.

L'apprentissage automatique non supervisé peut se diviser en deux classes :

- **Clustering** : l'objectif consiste à trouver des regroupements dans les données.
- **Association** : l'objectif consiste à identifier les règles qui permettront de définir de grands groupes de données.

Les principaux algorithmes du l'apprentissage automatique non supervisé sont les suivants : K-Means, clustering/regroupement hiérarchique et réduction de la dimensionnalité.[25]

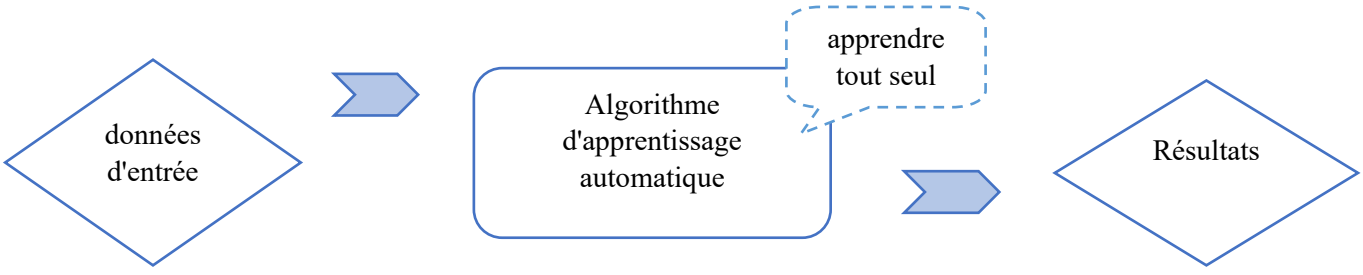


Figure 2-4 Apprentissage automatique non supervisé. [15]

	Apprentissage supervisé	Apprentissage non supervisé
Données d'entrée	Données connues en entrée	Données inconnues en entrée
Complexité informatique	Complexe	Moins complexe
Domaines d'activité	Classification et régression	Exploitation de règles de clustering et d'association
Précision	Produit des résultats précis	Génère des résultats modérés

Tableau 2.1 Différences l'apprentissage supervise et l'apprentissage non supervisé. [27]

2.6.3 L'apprentissage par renforcement

Cette approche implique un système d'agent qui interagit avec son environnement pour maximiser une récompense à long terme. Le modèle apprend ainsi à prendre des décisions en fonction de son état et de la récompense qu'il reçoit. Les exemples courants incluent les jeux, la robotique et la planification de trajectoire. [24] En 2013, c'était déjà un algorithme de machine learning par renforcement (Q-learning) qui s'était rendu célèbre en apprenant comment gagner dans six jeux vidéo Atari sans aucune intervention d'un programmeur.

L'apprentissage automatique par renforcement peut se diviser en deux classes :

- **Monte Carlo :** le programme reçoit ses récompenses à la fin de l'état « terminal ».
- **Machine learning par différence temporelle (TD)** : les récompenses sont évaluées et accordées à chaque étape.

Les principaux algorithmes du machine learning par renforcement sont les suivants : Q-learning, Deep Q Network (DQN) et SARSA (State-Action-Reward-State-Action).[25]

2.7 Les technologies d'intelligence artificielle open source

Les technologies de machine learning et d'intelligence artificielle (IA) modifient rapidement tous les domaines de notre société. Il semble que l'IA devient de plus en plus importante dans notre quotidien. En raison de ces progrès rapides, de nombreux talents et ressources sont consacrés à l'accélération de la croissance des technologies. Voici une liste des meilleures technologies d'intelligence artificielle open source.[28]

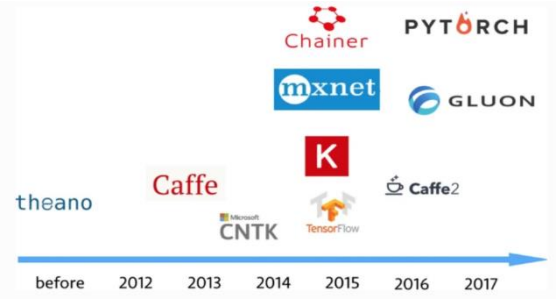
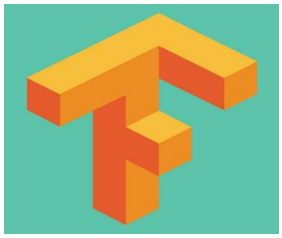


Figure 2-5 Chronologie des cadres d'apprentissage en profondeur.[29]

2.7.1 TensorFlow



Initialement publié en 2015, TensorFlow est un framework d'apprentissage machine open source facile à utiliser et à déployer. C'est l'un des cadres de machine learning les mieux entretenus et les plus utilisés.

- Produire par Google pour soutenir ses objectifs de recherche et de production, TensorFlow est maintenant très utilisé par plusieurs sociétés, notamment Dropbox, eBay, Intel, et Uber.TensorFlow peut être utilisé pour une variété de tâches d'apprentissage machine, y compris la classification, la régression, la segmentation d'image, la compréhension de la parole, la traduction automatique, la reconnaissance de caractères, l'analyse de texte et bien d'autres. Il est également possible de combiner plusieurs modèles pour créer des architectures de réseaux neuronaux complexes.

- TensorFlow fournit également des outils pour le prétraitement des données, la visualisation des modèles, la validation des modèles et la production de résultats, ainsi que des fonctionnalités pour l'exportation et l'importation de modèles. Il est également compatible avec une variété de langages de programmation, y compris Python, C++, Java, Rust, et plus récemment JavaScript, ce qui le rend accessible à un large public de développeurs.

- Concrètement, Le Framework vous permet de développer des réseaux de neurones (et même d'autres modèles de calcul) à l'aide de graphes de flux.

En résumé, TensorFlow est un cadre de développement de modèles de machine learning flexible et puissant qui peut être utilisé pour une grande variété de tâches d'apprentissage automatique.[28]

2.7.2 Keras

Paru en 2015, Keras est une bibliothèque de logiciels open source très populaire pour le machine learning. Elle a été conçue pour permettre aux développeurs de créer des modèles de machine learning de manière simple et rapide en offrant une interface de programmation conviviale et intuitive. Keras est codé en Python et peut être déployé sur d'autres technologies d'intelligence artificielle telles que TensorFlow, Microsoft Cognitive Toolkit (CNTK) et Theano. Keras est particulièrement apprécié pour sa convivialité, sa modularité et sa facilité d'extensibilité. Il est parfait si vous avez besoin d'un outil pour le machine learning permettant un prototypage simple et rapide, prenant en charge les réseaux convolutionnels et récurrents et fonctionnant de Cela rend Keras facile à apprendre et à utiliser. En tant qu'utilisateur de Keras, vous êtes plus productif, ce qui vous permet de tester plus d'idées et plus rapidement que vos concurrents, ce qui vous permet de gagner les compétitions de Machine Learning.

Cette simplicité d'utilisation n'impacte pas la flexibilité : Keras s'intègre avec les outils de Deep Learning (en particulier Tensorflow) à bas niveau, ce qui vous permet d'implémenter toutes les fonctionnalités que vous souhaiteriez utiliser. Keras s'intègre par exemple dans votre flux de travail sous TensorFlow, en tant que tf.keras . Keras est très largement utilisé par le monde industriel tout comme la communauté de chercheurs.

En novembre 2017, nous comptabilisions plus de 200 000 utilisateurs de Keras, ce qui en fait l'outil le plus utilisé dans le monde industriel et dans la recherche, après TensorFlow lui-même. (et Keras est habituellement utilisé conjointement avec TensrorFlow).

Vous interagissez déjà probablement sans le savoir avec des fonctionnalités conçues sous Keras, puisque la librairie est utilisée par Netflix, Uber, Yelp, Instacart, Zocdoc, Square et bien d'autres. Keras est surtout populaire chez les startups qui placent le Deep Learning au centre de leurs produits. Manière optimale à la fois sur les CPU (unités centrales) et les GPU (unités graphiques).

En résumé, Keras est une bibliothèque de logiciels open source très utile pour les développeurs de machine learning. Elle permet de créer rapidement des modèles de réseaux de neurones en utilisant une interface simple et conviviale, et est compatible avec plusieurs frameworks de deep learning.[28]

Keras permet de convertir facilement des modèles en des produits finis :

Les modèles peuvent être facilement déployés sur une plus large gamme de plateformes que les autres outils de Deep Learning :

- Sur iOS, via Apple CoreML (le support de Keras est officiellement fourni par Apple).
- Sur Android, avec le runtime TensorFlow pour android (exemple : l'app Not Hotdog)
- Dans le navigateur, avec des librairies supportant le GPU comme Keras.js et WebDNN.
- Sur Google Cloud, avec TensorFlow.
- En moteur d'application web Python (par exemple avec le framework Flask).
- Sur JVM, avec la fonctionnalité d'import de modèle DL4J fournie pa SkyMind.
- Sur Raspberry PI.[30]

2.7.3 Scikit Learn

Évoqué en 2007, Scikit-learn est largement utilisé dans l'industrie et la recherche pour diverses applications d'apprentissage automatique, telles que la classification d'images, la détection d'anomalies, la prédiction de valeurs, l'analyse de texte, etc. La bibliothèque est conçue sur trois autres projets open source NumPy, SciPy et Matplotlib, qui sont des bibliothèques scientifiques de base pour Python. Elle offre une interface simple et unifiée pour un large éventail de modèles

d'apprentissage automatique et fournit des outils pour la préparation et la manipulation des données, l'optimisation de modèle, l'évaluation de modèle, la validation croisée, etc.

Scikit-learn est également connu pour sa documentation détaillée, sa communauté active et ses fonctionnalités conviviales. La bibliothèque est en constante évolution avec de nouvelles fonctionnalités et améliorations ajoutées régulièrement. De plus, étant open source, il est facilement accessible pour les utilisateurs qui souhaitent contribuer ou personnaliser la bibliothèque en fonction de leurs besoins spécifiques.

En résumé, scikit-learn est une bibliothèque incontournable pour les développeurs Python qui travaillent dans le domaine de l'apprentissage automatique et de l'analyse de données.[28]

2.7.4 Theano

Initialement publié en 2007, Theano est un outil open source Python qui vous autoriser de concevoir facilement divers modèles d'apprentissage automatique. Étant donné qu'elle est l'une des plus anciennes bibliothèques, elle est considérée aujourd'hui comme un essentiel de l'industrie qui a inspiré les développements de deep learning. À la base, il vous permet de simplifier le processus de définition, d'optimisation et d'évaluation des expressions mathématiques. De plus, il est optimisé pour les GPU, offre une différenciation symbolique efficace et des fonctionnalités étendues de test de code. Theano est compatible avec de nombreuses autres bibliothèques Python utilisées pour l'apprentissage automatique, notamment NumPy, SciPy, Pandas et Scikit-learn. Il est également compatible avec les frameworks d'apprentissage profond tels que TensorFlow et Keras. Theano a été développé principalement par l'équipe de recherche de l'Université de Montréal sous la direction de Yoshua Bengio, l'un des pionniers de l'apprentissage profond. Il est conçu pour faciliter la création de réseaux de neurones profonds et leur entraînement en utilisant des techniques telles que la descente de gradient stochastique (SGD) et ses variantes. Theano offre une grande flexibilité dans la définition des modèles d'apprentissage automatique, permettant aux utilisateurs de définir des modèles personnalisés et de les entraîner sur des ensembles de données de grande taille. Il est également capable de tirer parti des GPU pour accélérer le processus d'entraînement.

Cependant, il convient de noter que Theano n'est plus activement développé et maintenu depuis 2017. Les développeurs recommandent désormais d'utiliser des bibliothèques plus récentes et plus performantes telles que TensorFlow ou PyTorch pour l'apprentissage profond.[28]

2.7.5 Caffe

Caffe est un framework d'apprentissage automatique open source qui a été publié pour la première fois en 2014. Il a été développé par Yangqing Jia lorsqu'il était étudiant au département d'informatique de l'Université de Californie, Berkeley, et depuis lors, il a été maintenu par une communauté de développeurs. Le framework open source est écrit en C ++ et est fourni avec une interface Python. Caffe est conçu pour faciliter la création et l'entraînement de réseaux de neurones convolutifs pour des tâches telles que la classification d'images, la détection d'objets et la segmentation sémantique. Il offre des performances rapides, en partie grâce à l'utilisation d'optimisations de bas niveau telles que BLAS et CUDA.

Caffe est également très flexible et modulaire, permettant aux utilisateurs de personnaliser facilement les architectures de réseau et de les combiner avec des modules de prétraitement et de post-traitement. Il est fourni avec une interface Python conviviale pour permettre aux utilisateurs de définir des modèles, d'entraîner des réseaux et de les déployer.

Cependant, il convient de noter que Caffe n'est plus activement développé et maintenu depuis 2018. Les développeurs recommandent désormais d'utiliser des bibliothèques plus récentes et plus performantes telles que TensorFlow ou PyTorch pour l'apprentissage profond.[28]

2.7.6 Torche

Torch est en effet une bibliothèque d'apprentissage automatique open source qui a été initialement publiée en 2002. Elle a été développée par l'équipe de recherche de l'Université de New York sous la direction de Ronan Collobert, et depuis lors, elle a été maintenue et développée par une communauté de développeurs. La structure open source vous offre une flexibilité et une rapidité optimisées lors du traitement de projets d'apprentissage automatique, sans entraîner de complexités inutiles. Torch est écrit en utilisant le langage de script Lua et utilise une implémentation en C sous-jacente pour améliorer les performances. Il offre une large gamme d'algorithmes pour l'apprentissage en profondeur, y compris des réseaux de neurones convolutifs, des réseaux de neurones récurrents et des réseaux de neurones de type perceptron multicouche.

Torch est également connu pour sa flexibilité et sa rapidité. Il offre des fonctionnalités telles que les baies N-dimensionnelles, les routines d'algèbre linéaire et les routines d'optimisation numérique, ainsi qu'une prise en charge efficace du processeur graphique pour accélérer l'entraînement des modèles d'apprentissage en profondeur. Torch est également disponible sur différentes plates-formes, y compris iOS et Android.

Cependant, il convient de noter que Torch n'est plus activement développé et maintenu depuis 2018.[28]

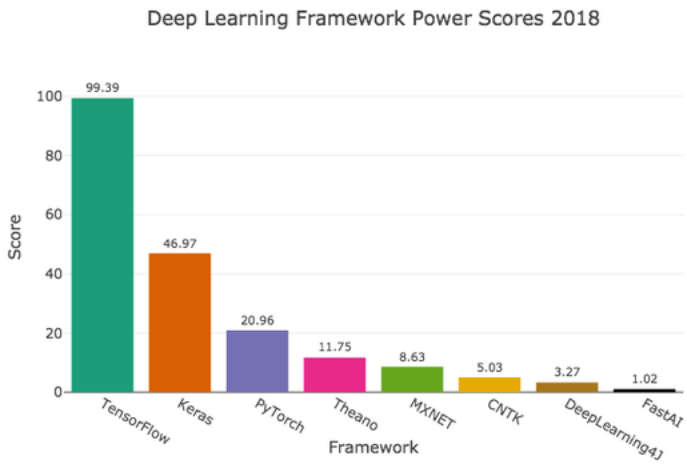


Figure 2-6 Scores de puissance du cadre d'apprentissage en profondeur 2018.[31]

2.8 Algorithmes de machine Learning

2.8.1 Les machines à vecteurs de support (SVM)

Les SVM sont des algorithmes populaires en raison de leur capacité à gérer des données de grandes dimensions et de leur simplicité d'utilisation. En ce qui concerne les hyperparamètres, il est vrai que le choix de la technique de régularisation et du noyau est important pour obtenir de bonnes performances. Le choix du noyau dépend notamment de la structure des données et de la nature du problème à résoudre. Par exemple, un noyau linéaire est souvent utilisé pour des problèmes de classification linéaire, tandis qu'un noyau gaussien est souvent utilisé pour des problèmes de classification non linéaire.[32] Le SVM est un algorithme de classification binaire à l'instar de la régression logistique. Si ces deux algorithmes permettent de séparer un ensemble d'éléments en deux classes. La machine à vecteur de support opte pour la plus nette séparation possible. Pour cela, les données sont séparées en plusieurs classes grâce à la « marge maximale ». C'est d'ailleurs pour cette raison qu'on la nomme Large Margins classifier. [33]

Les SVM fonctionnent en trouvant l'hyperplan qui maximise la marge entre les points de différentes classes dans un espace de dimensions élevées. Cela signifie que les SVM sont capables de travailler avec des données de grande dimension, ce qui les rend particulièrement adaptées pour la classification de données telles que les images, les séquences de texte et les données biomédicales. En plus de leur simplicité d'utilisation et de leur capacité à traiter des données de grande dimension, les SVM sont également appréciées pour leurs garanties théoriques. En effet, la théorie de Vapnik-Chervonenkis fournit des résultats mathématiques sur la performance des SVM et leur capacité à généraliser à de nouvelles données. En résumé, les SVM sont des modèles de machine learning supervisés populaires en raison de leur simplicité, de leur capacité à travailler avec des données de grande dimension, de leurs garanties théoriques et de leurs bons résultats en pratique. [34]

En fait, l'objectif principal est de séparer l'ensemble de données de la meilleure façon possible. La distance entre les deux points les plus proches est connue sous le nom de marge. L'objectif est de sélectionner un hyperplan avec la marge maximale possible entre les vecteurs de support dans l'ensemble de données. Pour chercher, l'hyperplan marginal maximal, le SVM commence par générer des hyperplans qui séparent au mieux les classes. Ensuite, il sélectionne l'hyperplan droit avec la ségrégation maximale des points de données les plus proches, comme indiqué dans la figure ci-dessous.[35]

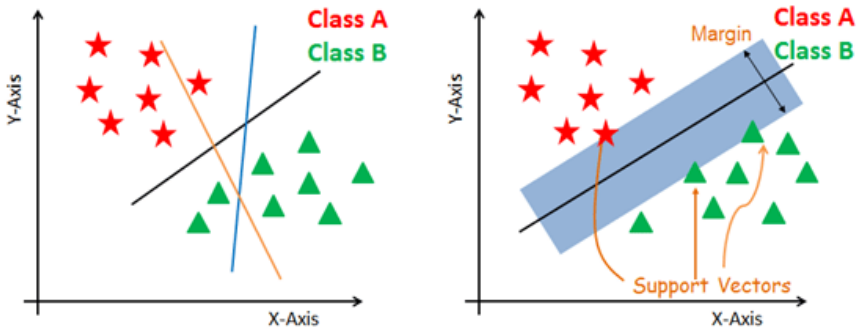


Figure 2-7 Support Vector Machine (SVM). [35]

2.8.2 Régression linéaire

La régression linéaire est l'une des techniques de modélisation et de prédiction les plus anciennes et les plus simples en statistique. Elle repose sur l'hypothèse que la relation entre la variable dépendante et les variables indépendantes peut être modélisée par une fonction linéaire.

Un modèle de régression linéaire simple est de la forme :

$$Y = \beta 0 + \beta 1X1 + \beta 2X2 + \beta 3X3 + \beta 4X4 + \cdots + \varepsilon$$

Où

Y est la variable dépendante ou le résultat que nous essayons de prédire.

β0 et β1 sont les coefficients (ordonnée à l'origine et pente).

X1.X2.X3 est les variables indépendantes utilisées pour prédire le résultat (variable explicative).

ε est une erreur aléatoire.

Un modèle de régression linéaire a deux composantes : une partie déterministe (i.e. b1X1 + b2X2 + ...) et une partie aléatoire (i.e. l'erreur, ε). On peut considérer ces deux composants comme le signal et le bruit dans le modèle. Si On n'a qu'une seule variable d'entrée X, le modèle de régression est la meilleure ligne qui correspond aux données. La figure 2.8 montre un exemple de modèle de régression linéaire simple. Avec deux variables d'entrée, la régression linéaire est le meilleur plan qui s'adapte à un ensemble de points de données dans un espace 3D.

Les coefficients des variables (i.e. β1, β2, β3, etc.) sont les pentes partielles de chaque variable. Si on maintient toutes les autres variables constantes, le résultat Y augmentera de β1 lorsque la variable X1 augmentera de 1. C'est pourquoi les économistes utilisent généralement l'expression « ceteris paribus » ou « toutes choses étant égales par ailleurs » pour décrire l'effet d'une variable indépendante sur un résultat donné.[36]

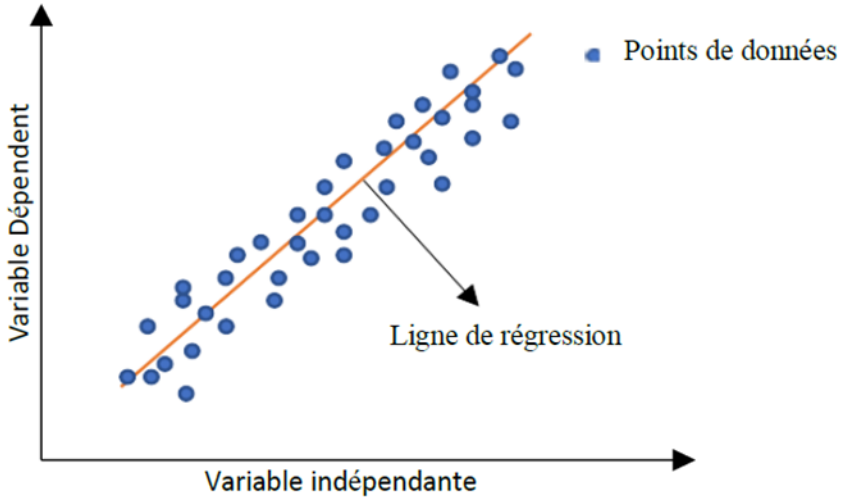


Figure 2-8 Un modèle de régression linéaire simple.[37]

Malgré sa simplicité, la régression linéaire reste une technique précieuse et largement utilisée dans de nombreux domaines pour la prédiction, l'inférence et l'exploration de relations linéaires entre les variables.

La régression linéaire utilise les méthodes des moindres carrés ou de descente de gradient pour déterminer les coefficients optimaux du modèle à partir d'un ensemble de données donné. La méthode des moindres carrés minimise la somme des erreurs quadratiques entre les valeurs prédites par le modèle et les valeurs réelles de chaque observation dans les données d'apprentissage. En ajustant les coefficients à chaque itération, la descente de gradient cherche à réduire la somme des erreurs entre les valeurs prédites par le modèle ajusté et les valeurs réelles des données d'apprentissage. Par le biais de plusieurs itérations, elle converge vers un minimum local en se déplaçant dans la direction opposée au gradient négatif.[36]

2.8.3 Algorithme Ridge

L'algorithme Ridge, également connu sous le nom de régression Ridge, est une technique de régression linéaire régularisée. Il est utilisé pour traiter le problème de la régression lorsque les données d'entraînement sont sujettes à une multicollinéarité élevée, c'est-à-dire lorsque certaines variables indépendantes sont fortement corrélées entre elles. L'objectif de la régression Ridge est de trouver les coefficients du modèle linéaire qui minimisent à la fois l'erreur de prédiction et la somme des carrés des coefficients (aussi appelée pénalité de régularisation). Cette pénalité de régularisation est contrôlée par un hyperparamètre appelé alpha (λ), qui détermine le niveau de régularisation appliqué.

Le principe de base de l'algorithme Ridge est d'ajouter un terme de régularisation à la fonction de coût de la régression linéaire. Ce terme de régularisation est proportionnel à la somme des carrés des coefficients du modèle. En ajoutant ce terme, les coefficients du modèle sont contraints à rester petits, ce qui réduit l'impact des variables fortement corrélées et aide à atténuer les effets de la multicollinéarité.

Dans le cas de la régression des moindres carrés, la fonction de risque est :

$$\sum_{i=1}^{n}(Yi - X\beta|_{¡})^2 = \sum_{i=1}^{n}(Yi - \beta 0 - \sum_{j=1}^{q}\beta jxij)^2$$

La fonction de coût pour Ridge régression :

$$Min\ (||\ y{-}x(\theta)\ ||^2{+}\ \lambda||\theta||^2)$$

Lambda est le terme de pénalité. λ donné ici est noté par un paramètre alpha dans la fonction de Ridge. Ainsi, en changeant les valeurs d'alpha, nous contrôlons le terme de pénalité. Plus les valeurs d'alpha sont élevées, plus la pénalité est grande et donc l'amplitude des coefficients est réduite.

- Il réduit les paramètres. Par conséquent, il est utilisé pour empêcher la multicolinéarité.

- Il réduit la complexité du modèle par le retrait du coefficient.

Pour tout type de modèles d'apprentissage automatique de régression, l'équation de régression habituelle constitue la base qui s'écrit comme suit :

$$Y{=}X\ B{+}e$$

Où *Y* est la variable dépendante, *X* représente les variables indépendantes, *B* est les coefficients de régression à estimer et *e* représente les erreurs sont des résidus.

Une fois que nous ajoutons la fonction lambda à cette équation, la variance qui n'est pas évaluée par le modèle général est considérée.

Les étapes principales de l'algorithme Ridge :

1. **Préparation des données :** Tout d'abord, les données d'entraînement sont préparées en séparant les variables indépendantes (caractéristiques) de la variable cible à prédire.
2. **Normalisation des données** : Pour garantir des résultats cohérents, il est souvent recommandé de normaliser les caractéristiques en soustrayant la moyenne et en divisant par l'écart-type. Cette étape est importante pour s'assurer que toutes les caractéristiques sont sur une échelle comparable.
3. **Ajout d'un terme de régularisation :** L'algorithme Ridge ajoute un terme de régularisation à la fonction de coût de la régression linéaire. Ce terme est proportionnel à la somme des

carrés des coefficients du modèle, pondéré par un hyperparamètre alpha (λ). Plus la valeur d'alpha est grande, plus la pénalité de régularisation est forte.

4. **Minimisation de la fonction de coût :** L'algorithme Ridge utilise des techniques d'optimisation numérique pour trouver les coefficients qui minimisent la fonction de coût régularisée. La méthode des moindres carrés ou la descente de gradient peuvent être utilisées pour cette minimisation.
5. **Choix de l'hyperparamètre alpha :** Le choix approprié de l'hyperparamètre alpha est crucial dans l'algorithme Ridge. Une valeur trop élevée d'alpha peut entraîner un sous-ajustement (underfitting), tandis qu'une valeur trop faible peut ne pas suffisamment régulariser le modèle. Des techniques telles que la validation croisée peuvent être utilisées pour sélectionner la meilleure valeur d'alpha.
6. **Évaluation du modèle :** Une fois que les coefficients optimaux ont été trouvés, le modèle peut être évalué sur un ensemble de données de test pour estimer ses performances de prédiction. Les métriques d'évaluation courantes pour la régression linéaire incluent l'erreur quadratique moyenne (RMSE) et le coefficient de détermination (R^2).

L'algorithme Ridge permet de trouver un compromis entre l'ajustement aux données d'entraînement et la complexité du modèle. En ajoutant une pénalité de régularisation, il contribue à réduire les effets de la multicollinéarité et à améliorer la généralisation du modèle sur de nouvelles données.[38]

2.8.4 Algorithme Lasso

L'algorithme Lasso (Least Absolute Shrinkage and Selection Operator) est une méthode de régression pénalisée utilisée en apprentissage automatique et en statistiques. Il est couramment utilisé pour la sélection de variables et la régularisation dans les modèles de régression linéaire. Le Lasso vise à estimer les coefficients d'un modèle linéaire tout en introduisant une pénalité qui pousse certains coefficients vers zéro, ce qui permet de sélectionner automatiquement les variables les plus pertinentes pour la prédiction. Cela peut être utile lorsque vous avez un grand nombre de variables explicatives et que vous souhaitez identifier celles qui ont le plus d'impact sur la variable cible.

L'idée clé du Lasso est d'ajouter un terme de pénalité L1 à la fonction de coût du modèle de régression linéaire. Ce terme de pénalité est proportionnel à la somme des valeurs absolues des coefficients de régression. Ainsi, lors de l'optimisation du modèle, le Lasso favorise des solutions où de nombreux coefficients sont exactement zéro, ce qui permet une sélection automatique des variables. La force de pénalité du Lasso est contrôlée par un paramètre appelé alpha. Plus la valeur d'alpha est élevée, plus la pénalité est forte et plus de coefficients seront poussés vers zéro. En ajustant la valeur d'alpha, vous pouvez ajuster le niveau de régularisation et contrôler la quantité de variables sélectionnées. L'algorithme Lasso peut être résolu à l'aide de différentes techniques d'optimisation, telles que la descente de gradient coordonnée (coordinate descent) ou la méthode du lagrangien augmenté (augmented Lagrangian method). La régression lasso correspond à un choix de pénalité particulier.

A la pénalité $\lambda\sum_{i=1}^{q}\beta^2$j de la régression ridge, on substitue la pénalité :

$$\lambda\sum_{i=1}^{q}|\beta j\,|$$

La première pénalité fait intervenir la norme L1, llβlll = $\sum_{i=1}^{q}|\beta j\,|$ et est parfois appelée pénalité L1. La seconde pénalité fait intervenir la norme L2, llβlll = $\sum_{i=1}^{q}\beta^2 j$ et est appelée pénalité L2. Les coefficients de régression lasso sont obtenus en minimisant une fonction de risque de la forme :

$$\sum_{i=1}^{n}(Yi - X\,\beta|_{\text{i}})^2 + \lambda\sum_{j=1}^{q}|\beta j\,|, \lambda \geq 0$$

Comme la régression ridge, la régression lasso génère des coefficients de régression qui tendent vers 0 mais peuvent également être égaux à 0 pour une 3 2 Régression Lasso valeur de λ assez grande.[38]

2.8.5 L'arbre de décision

L'arbre de décision est une représentation basée sur des règles logiques ou de production qui est construite automatiquement à partir d'un ensemble de données. La construction de l'arbre de décision se fait en subdivisant progressivement l'ensemble de données en sous-ensembles de plus en plus spécifiques à l'aide des descripteurs (caractéristiques) disponibles.

Un arbre de décision est un algorithme d'apprentissage supervisé qui permet de prendre des décisions en utilisant une structure arborescente. Il est couramment utilisé pour la classification, où l'objectif est de prédire une classe ou une catégorie à partir des caractéristiques d'un exemple donné, ainsi que pour la régression, où l'objectif est d'estimer une valeur numérique. La structure d'un arbre de décision est composée d'un nœud racine qui représente la caractéristique la plus importante pour la division initiale, puis de branches qui représentent les différentes valeurs possibles de cette caractéristique. À partir du nœud racine, l'arbre se divise en nœuds internes, qui représentent d'autres caractéristiques et leurs valeurs, et en nœuds feuilles, qui correspondent aux prédictions finales. L'algorithme estime la probabilité qu'une observation se trouve dans le noeud t en utilisant la formule :

$$P(T) = \sum_{j\in T} w_j$$

Où :

w_j: Poids des observations,

T : Ensemble de tous les indices d'observation du nœud t

n : Égale au nombre d'observations.

Les arbres de décision sont des algorithmes de classification de données reposant sur un apprentissage supervisé. Comme son nom l'indique, la structure de ces algorithmes ressemble à des arbres constitués de nœuds, de branches et de feuilles. La construction de ces arbres est réalisée à l'aide d'une base de données brute (vecteur des caractéristiques et classes) et de lois qui permettent de déterminer les variables discriminantes pour la classification efficace des futures données. Chacun des nœuds constituant l'arbre représente une règle de classification préalablement déterminée de manière récursive.

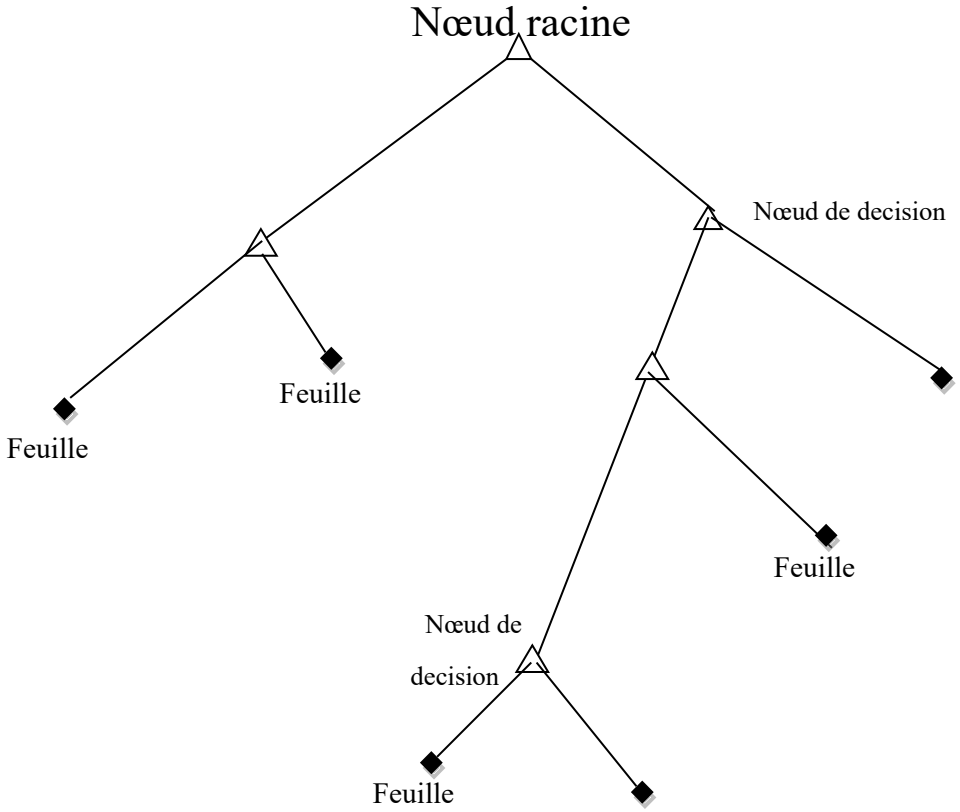


Figure 2-9 Exemple d'illustration d'un arbre de décision.[39]

En fouille de données, on distingue principalement deux types d'arbres de décision :

- **Les arbres de classification (Classification Trees) :**

Ils sont utilisés pour prédire la classe à laquelle une variable cible appartient. Dans ce cas, la prédiction résulte en une étiquette de classe. L'arbre est construit en utilisant des tests sur les variables indépendantes pour diviser les données en sous-groupes homogènes. Chaque feuille de l'arbre représente une classe prédite et lorsqu'une nouvelle observation est présentée à l'arbre, elle suit les branches correspondantes aux tests jusqu'à atteindre une feuille, où la classe prédite est assignée.

- **Les arbres de régression (Regression Trees) :**

Ils sont utilisés pour prédire une quantité réelle, telle que le prix d'une maison ou la durée de séjour d'un patient dans un hôpital. La prédiction résulte donc en une valeur numérique. De manière similaire aux arbres de classification, les arbres de régression utilisent des tests sur les variables indépendantes pour diviser les données en sous-groupes. Les feuilles de l'arbre contiennent les valeurs numériques prédites. Lorsqu'une nouvelle observation est présentée à l'arbre, elle suit les branches en fonction des tests effectués jusqu'à atteindre une feuille, où la valeur numérique prédite est attribuée.[39]

2.8.6 K-plus proche voisin (KNN)

Le k-plus proche voisin (KNN), est un algorithme simple d'apprentissage automatique supervisé utilisé pour la classification et la régression. Il est souvent utilisé dans des problèmes où les données sont étiquetées et où il faut prédire une étiquette pour de nouvelles instances non étiquetées. L'algorithme recherche les k instances les plus similaires à cette observation. Ensuite, il détermine la classe majoritaire parmi ces k plus proches voisins et attribue cette classe à l'observation en question. Elle est considérée comme l'une des techniques les plus simples pour classer de nouvelles observations.

Le fonctionnement de l'algorithme k-NN est le suivant :

1. L'ensemble de données d'entraînement contient des exemples avec leurs étiquettes de classe correspondantes.
2. Lorsqu'une nouvelle observation doit être classifiée, l'algorithme calcule la distance entre cette observation et tous les exemples du jeu de données d'entraînement. La distance peut être calculée à l'aide de différentes mesures, telles que la distance euclidienne.
3. Les k exemples les plus proches de la nouvelle observation sont sélectionnés en fonction de leur distance.
4. La classe majoritaire parmi les k plus proches voisins est attribuée à la nouvelle observation, ce qui la classe dans cette classe.

Il est important de noter que le choix de la valeur de k est un paramètre important de l'algorithme k-NN. Une valeur trop petite de k peut rendre le modèle sensible au bruit ou aux points aberrants, tandis qu'une valeur trop grande peut conduire à une classification moins précise.

Bien que le temps d'apprentissage de l'algorithme k-NN soit court, le temps de requête réel (et l'espace de stockage) peut être plus long que celui des autres modèles. Cela est particulièrement vrai lorsque le nombre de points de données augmente, car toutes les données d'entraînement doivent être conservées, mais pas seulement l'algorithme.
Bien que l'algorithme des k-plus proches voisins (k-NN) soit simple et facile à comprendre, il présente certains inconvénients :

- Sensible à la dimensionnalité : L'efficacité de k-NN diminue lorsque le nombre de dimensions des données augmente. Cela est dû au phénomène connu sous le nom de "malédiction de la dimensionnalité", où les distances entre les points de données deviennent moins significatives dans des espaces de grande dimension, rendant la classification moins fiable.

- Sensible aux valeurs aberrantes : k-NN peut être sensible aux valeurs aberrantes (outliers) car elles peuvent influencer les calculs de distance et conduire à des classifications erronées. Une observation aberrante peut avoir un impact disproportionné sur la classification si elle se trouve parmi les k plus proches voisins.
- Besoin de définir une valeur de k : Le choix d'une valeur appropriée pour k est crucial dans k-NN. Une valeur de k trop petite peut conduire à une classification bruitée et instable, tandis qu'une valeur de k trop grande peut entraîner une perte de précision et une fusion des frontières de décision.
- Calcul intensif des distances : L'algorithme k-NN nécessite de calculer les distances entre la nouvelle observation et tous les exemples d'apprentissage, ce qui peut être coûteux en termes de temps de calcul, en particulier pour les ensembles de données volumineux. De plus, les calculs de distance peuvent devenir plus complexes dans des espaces de grande dimension.

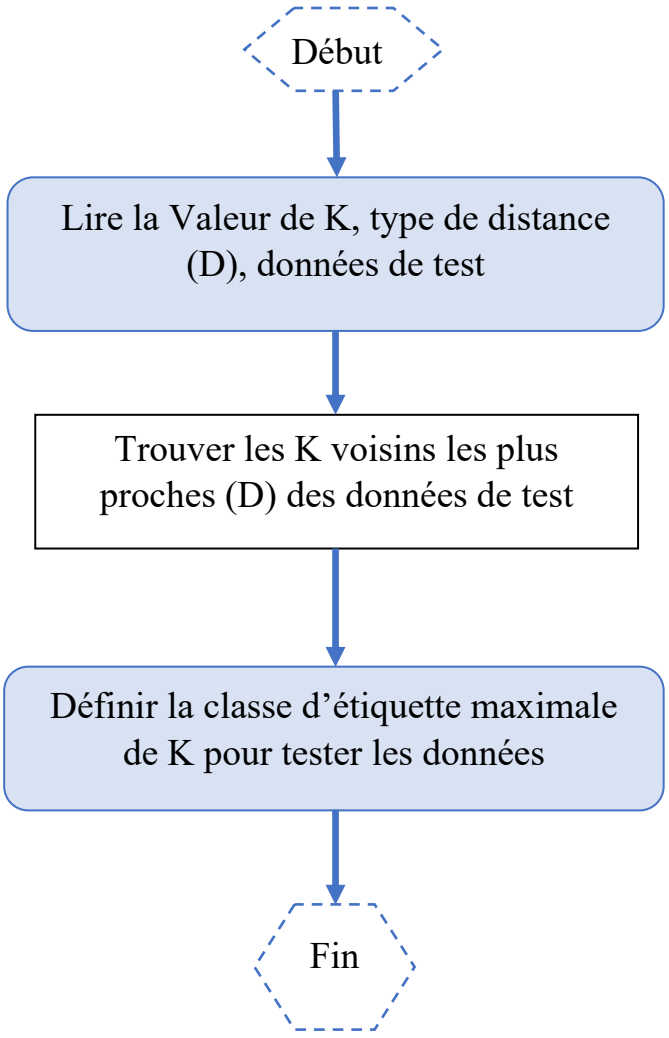


Figure 2-10 Organigramme de l'algorithme KNN.[39]

L'algorithme k-NN calcule la distance entre les points de données. Pour cela, nous utilisons la formule de la distance euclidienne [39] :

$$d(p, q) = d(q, p) = \sqrt{(q1 - p1)^2 + (q2 - p2)^2} + \cdots + (qn - pn)$$

$$\boldsymbol{d(p,q) = d(q,p) = \sqrt{\sum_{i=1}^{n}(qi - pi)^2}}$$

2.8.7 NAÏVE BAYES

Les algorithmes Naïve Bayes sont couramment utilisés lorsque les données ne sont pas complexes et que la tâche de classification est relativement simple. Ils sont particulièrement adaptés aux problèmes de classification textuelle, tels que la catégorisation d'e-mails en spam ou non-spam, la classification de documents, etc. Le principe de base de Naïve Bayes est de calculer la probabilité qu'un exemple donné appartienne à chaque classe possible, puis de classer l'exemple dans la classe ayant la probabilité la plus élevée. Pour ce faire, l'algorithme utilise les probabilités conditionnelles et l'hypothèse d'indépendance conditionnelle, qui suppose que les caractéristiques (ou variables prédictives) sont indépendantes les unes des autres.

L'hypothèse d'indépendance conditionnelle permet de simplifier le calcul des probabilités conditionnelles en supposant que chaque caractéristique contribue indépendamment à la probabilité d'appartenance à une classe donnée. Cela permet de réduire la complexité du modèle et d'accélérer l'apprentissage et la prédiction.

L'algorithme Naïve Bayes nécessite un ensemble d'exemples d'apprentissage étiquetés pour estimer les probabilités conditionnelles. À partir de ces exemples, il calcule les probabilités a priori des classes et les probabilités conditionnelles des caractéristiques pour chaque classe. Lorsqu'un nouvel exemple est présenté, l'algorithme utilise ces probabilités pour estimer la probabilité d'appartenance à chaque classe et classe l'exemple dans la classe avec la probabilité la plus élevée. Le principal avantage des algorithmes Naïve Bayes réside dans leur simplicité et leur efficacité. Ils sont relativement rapides à entraîner et à appliquer sur de nouveaux exemples, ce qui les rend attrayants pour des tâches avec une quantité limitée de données disponibles pour l'apprentissage et ce qui en fait un choix populaire lorsque les données ne sont pas complexes et que la tâche de classification est relativement simple.[39]

$$P(A/B) = \frac{P(B/A)\,.\,P(A)}{P(B)}$$

2.8.8 Régression polynomiale

La régression polynomiale est une technique d'analyse statistique utilisée pour modéliser des relations non linéaires entre une variable dépendante et une ou plusieurs variables indépendantes.

Elle est une extension de la régression linéaire, qui suppose une relation linéaire entre les variables. Les fonctions polynomiales plus complexes pour généraliser sur des données non linéaires. En réalité, le seul changement apporté est basé sur notre fonction d'hypothèse puisque la fonction d'erreur et les deux algorithmes restent exactement les mêmes. Il suffit donc d'employer une fonction d'hypothèse polynomiale :

Dans cette expression, d représente le degré maximum de notre fonction.

Où $h\theta(x)$ est la valeur prédite de la variable dépendante y pour une valeur donnée de la variable indépendante x, θ_0, θ_1, θ_2, ..., θ_n sont les coefficients du polynôme et n est le degré du polynôme.

L'idée de base est d'inclure des termes polynomiaux de degrés supérieurs dans la fonction d'hypothèse pour capturer des relations non linéaires entre les variables. Par exemple, pour un polynôme de degré 2, la fonction d'hypothèse devient :

$$h\theta(x) = \theta_0 + \theta_1 x + \theta_2 x^2$$

Il est important de noter que la régression polynomiale peut être sujette au surajustement (overfitting) si le degré du polynôme est trop élevé par rapport à la taille des données. Dans de tels cas, le modèle peut être trop complexe et avoir du mal à généraliser correctement sur de nouvelles données. La sélection du degré du polynôme est une décision importante dans la régression polynomiale et peut être basée sur l'analyse des données, la validation croisée ou d'autres méthodes de sélection de modèle.[40]

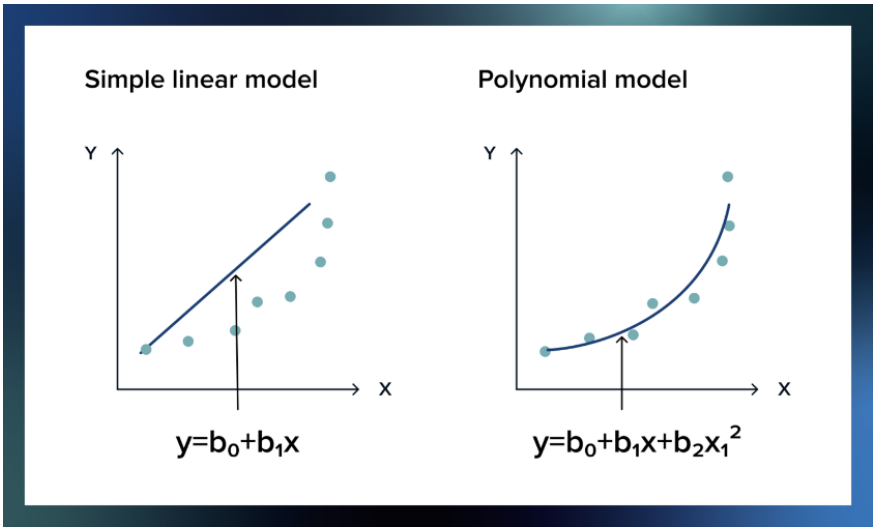


Figure 2-11 Exemple de régression polynomiale et linéaire.[41]

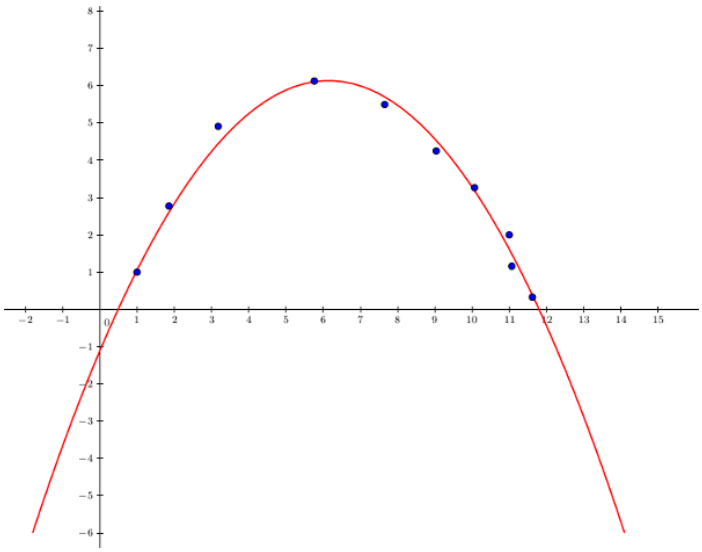


Figure 2-12 Régression linéaire et polynomiale.[40]

En résumé, la régression polynomiale utilise une fonction d'hypothèse polynomiale pour modéliser des relations non linéaires entre les variables. Elle permet de généraliser la régression linéaire pour des données non linéaires en ajoutant des termes polynomiaux de degrés supérieurs.

2.8.9 Régression logistique

La régression logistique est une technique d'apprentissage supervisé utilisée pour prédire une variable binaire ou catégorielle à partir de variables indépendantes. Contrairement à la régression linéaire qui prédit des valeurs continues, la régression logistique est utilisée pour des tâches de classification. La régression logistique est en effet utilisée dans des tâches de classification et est appropriée lorsque la variable dépendante est dichotomique (deux résultats possibles) ou catégorielle. L'objectif de la régression logistique est de trouver le meilleur modèle approprié pour décrire les relations au sein de la caractéristique dichotomique d'intérêt et d'un ensemble d'attributs autonomes.

Les types de régressions logistiques mentionnés :

- Régression logistique binaire : Il s'agit du cas le plus courant où la variable dépendante est binaire, c'est-à-dire qu'elle ne prend que deux valeurs possibles, généralement codées comme 0 et 1. La régression logistique binaire est utilisée pour modéliser la probabilité d'appartenir à la classe 1 en fonction des variables indépendantes.
- Régression logistique multinomiale : Aussi appelée régression logistique polytomique, elle est utilisée lorsque la variable dépendante a plus de deux catégories mutuellement exclusives. Par exemple, si la variable dépendante représente le type de fleur (rose, tulipe,

marguerite), la régression logistique multinomiale peut prédire la probabilité d'appartenir à chaque type de fleur en fonction des variables indépendantes.

- Régression logistique ordonnée : Ce type de régression logistique est utilisé lorsque la variable dépendante est ordonnée ou catégorielle avec un ordre naturel. Par exemple, si la variable dépendante représente le niveau de satisfaction (faible, moyen, élevé), la régression logistique ordonnée permet de modéliser la probabilité d'appartenir à chaque niveau de satisfaction en fonction des variables indépendantes.[42]

2.8.10 Random Forest

Random Forest est un algorithme d'apprentissage supervisé qui construit un ensemble d'arbres de décision en utilisant des échantillons aléatoires et des sous-ensembles aléatoires de variables. Comme son nom l'indique (forêt aléatoire), un RF crée un ensemble (une forêt) avec plusieurs arbres de décision aléatoires. Il agrège ensuite les prédictions des arbres individuels pour fournir une prédiction finale. C'est un algorithme puissant pour la classification et la régression, offrant des performances élevées et une réduction du surajustement.

L'avantage majeur de Random Forest réside dans sa capacité à réduire le surajustement (overfitting) par rapport à un arbre de décision individuel. En combinant les prédictions de plusieurs arbres, il réduit la variance et améliore la généralisation du modèle. Random Forest est également robuste face à des jeux de données contenant un grand nombre de variables et peut fournir des estimations de l'importance des variables pour la prédiction.

Le but de la méthode RF est d'obtenir des résultats plus efficaces avec plus d'un décideur comme dans d'autres méthodes. La différence de cette méthode par rapport aux autres méthodes est que les variables sont sélectionnées au hasard lorsque les branches se ramifient.[43]

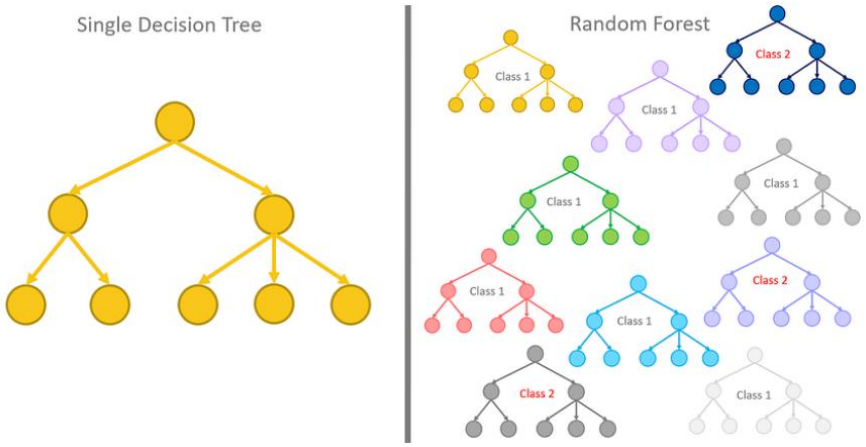


Figure 2-13 L'algorithme de forêt aléatoire.[44]

2.8.11 Réseaux de neurones artificiels

Les réseaux de neurones artificiels (Artificial Neural Networks), également appelés réseaux neuronaux ou réseaux de neurones profonds, Il a été introduit par Frank Rosenblatt en 1957, mais son utilisation réelle n'a été largement répandue qu'à partir de 1982, après des améliorations significatives, sont des modèles d'apprentissage automatique inspirés par le fonctionnement du cerveau humain. Ils sont utilisés pour des tâches de classification, de régression, de reconnaissance de motifs, de traitement du langage naturel et bien d'autres domaines. Un réseau de neurones artificiels est composé de plusieurs couches de neurones interconnectés. Chaque neurone est une unité de calcul qui reçoit des entrées pondérées, les combine à l'aide d'une fonction d'activation et produit une sortie. Les couches intermédiaires, appelées couches cachées, permettent d'apprendre des représentations hiérarchiques des données.

L'apprentissage d'un réseau de neurones se fait généralement par rétropropagation du gradient. Cela consiste à ajuster les poids des connexions entre les neurones pour minimiser une fonction de perte qui mesure l'écart entre les prédictions du réseau et les valeurs réelles. L'algorithme de descente de gradient est utilisé pour mettre à jour les poids de manière itérative jusqu'à ce que le modèle atteigne une certaine convergence.

Il existe différentes architectures de réseaux de neurones, telles que les réseaux de neurones à propagation directe (feedforward), les réseaux de neurones récurrents (RNN) pour le traitement de séquences, et les réseaux de neurones convolutifs (CNN) pour l'analyse d'images. Les réseaux de neurones artificiels ont été utilisés avec succès dans de nombreux domaines, notamment la vision par ordinateur, la reconnaissance vocale, la traduction automatique et les systèmes de recommandation. Leur puissance et leur flexibilité en font des outils puissants pour résoudre une grande variété de problèmes d'apprentissage automatique.

Domaines d'application des réseaux de neurones artificiels Aujourd'hui, les réseaux de neurones artificiels ont de nombreuses applications dans des secteurs très variés :

- **Traitement d'images :** reconnaissance de caractères et de signatures, compression d'images, reconnaissance de forme, cryptage, classification, etc.
- **Traitement du signal :** filtrage, classification, identification de source, traitement de la parole…etc.
- **Contrôle :** commande de processus, diagnostic, contrôle qualité, asservissement des robots, systèmes de guidage automatique des automobiles et des avions…etc.
- **Défense :** guidage des missiles, suivi de cible, reconnaissance du visage, radar, sonar, lidar, compression de données, suppression du bruit…etc.
- **Optimisation :** planification, allocation de ressource, gestion et finances, etc.
- **Simulation : simulation** du vol, simulation de boîte noire, prévision météorologique, recopie de modèle…etc.

La fonction d'activation sigmoïde est couramment utilisée dans les réseaux de neurones est définie comme suit :

$$f(x) = \frac{1}{1 + e^{-x}}$$

Les réseaux de neurones artificiels sont excellents pour modéliser des données non linéaires avec un grand nombre de fonctionnalités d'entrée. Lorsqu'ils sont utilisés correctement, les ANN peuvent résoudre des problèmes qui sont trop difficiles à résoudre avec un algorithme simple. Cependant, les réseaux de neurones sont coûteux en calcul, il est difficile de comprendre comment un réseau de neurones artificiels a atteint une solution.[45]

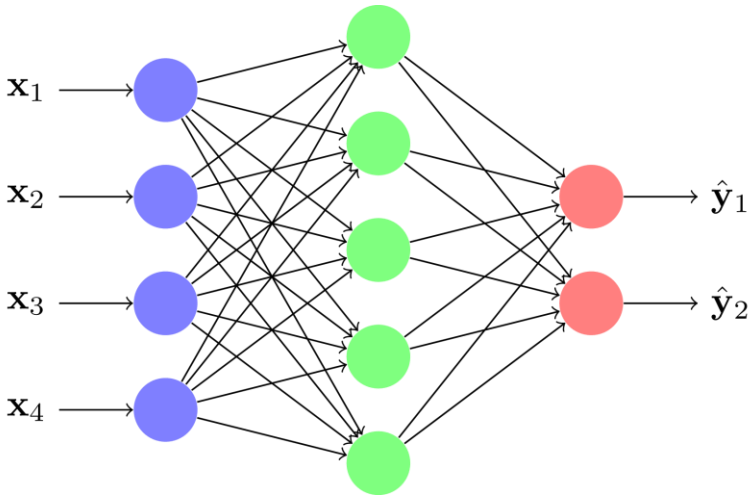


Figure 2-14 Un réseau de neurones.[46]

2.8.12 Le perceptron multicouche

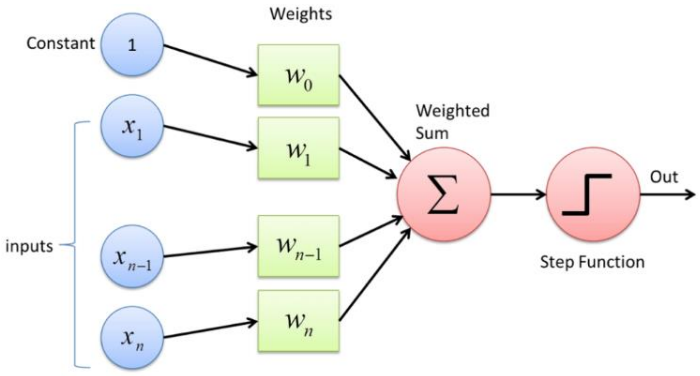


Figure 2-15 Fonctionnement du perceptron multicouche.[47]

Les réseaux de neurones artificiels, également appelés réseaux de neurones, sont des modèles mathématiques inspirés par le fonctionnement des neurones dans le cerveau humain. Ils sont utilisés dans de nombreux domaines de l'intelligence artificielle, tels que la vision par ordinateur,

le traitement du langage naturel, la reconnaissance de formes, etc. Le perceptron est un cas particulier de réseau de neurones artificiels, qui est simple mais limité dans sa capacité à modéliser des relations non linéaires. Les réseaux de neurones plus complexes, tels que les réseaux de neurones profonds, permettent de résoudre des problèmes plus complexes en utilisant des architectures avec plusieurs couches cachées.

Le perceptron est une forme simple de réseau de neurones, souvent considérée comme le bloc de construction de base des réseaux de neurones plus complexes. Il est composé d'une seule couche de neurones, appelée couche d'entrée, et d'une couche de sortie. Chaque neurone dans la couche d'entrée est connecté à tous les neurones dans la couche de sortie. Chaque connexion entre un neurone d'entrée et un neurone de sortie est associée à un poids. Le perceptron effectue un calcul pondéré des entrées en multipliant chaque entrée par son poids correspondant, puis en les sommant. Cette somme pondérée est ensuite passée à une fonction d'activation, qui produit la sortie du neurone. La fonction d'activation est généralement une fonction non linéaire, telle que la fonction sigmoïde ou la fonction de seuil. La sortie du perceptron peut être interprétée comme une prédiction ou une classification. Par exemple, dans un problème de classification binaire, la sortie du perceptron peut être 0 ou 1, représentant les deux classes possibles. Le processus d'apprentissage du perceptron consiste à ajuster les poids des connexions de manière itérative afin de minimiser une fonction d'erreur. Un algorithme couramment utilisé pour l'apprentissage du perceptron est l'algorithme de descente de gradient, qui ajuste les poids dans la direction qui minimise l'erreur.

Il est important de noter que le perceptron est un modèle linéaire, ce qui signifie qu'il ne peut apprendre que des relations linéaires entre les variables d'entrée et de sortie. Pour modéliser des relations plus complexes, des réseaux de neurones multicouches, tels que les réseaux de neurones profonds, sont utilisés. Ces réseaux comportent plusieurs couches cachées entre la couche d'entrée et la couche de sortie, permettant ainsi d'apprendre des représentations hiérarchiques des données.[39]

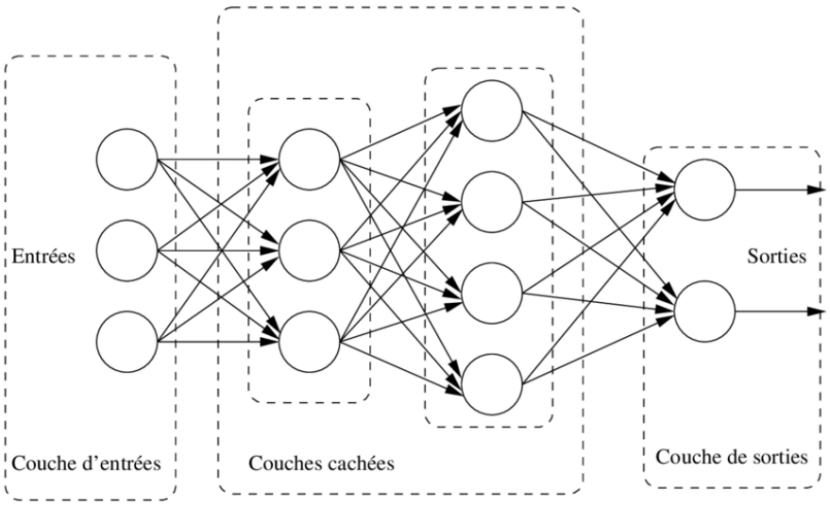


Figure 2-16 perceptrons multicouches.[48]

2.9 Etat de l'art sur le ML pour la maintenance prédictive

Récemment, la maintenance prédictive est un domaine d'application important de l'apprentissage automatique. Qui vise à prédire les pannes d'équipements avant qu'elles ne se produisent, permettant ainsi aux entreprises de planifier la maintenance à l'avance et d'éviter des temps d'arrêt coûteux. Les chercheurs ont récemment proposé de nombreuses approches basées sur les données pour résoudre les problèmes de maintenance prédictive. Les méthodes les plus couramment utilisées comprennent la détection d'anomalies, la classification de défauts, l'estimation de la durée de vie restante et la prise de décision pour les stratégies de maintenance. Les chercheurs ont lancé beaucoup de travaux de recherche afin d'appliquer les techniques de l'apprentissage automatique à la MI. Par ex, à proposer une nouvelle approche de pronostic basé sur le modèle cloud computing et le principe de multitenancy afin de présenter le pronostic en tant que service. Son approche fournit une solution de pronostic efficace à la demande d'un client tout en assurant une meilleure qualité du service. Comme exemple on retrouve l'implémentation et le test sur des données de moteurs d'avions de la NASA trois méthodes de pronostic guidé par les données (réseau de neurones artificiels, système neuro-flou et réseau bayésien) afin de tester l'efficacité de sa solution comparer les méthodes implémentées.

Il est intéressant de noter que l'application des techniques de l'apprentissage automatique à la maintenance industrielle est un domaine de recherche en plein essor, avec de nombreux travaux de recherche et projets en cours. L'utilisation de modèles de prédiction basés sur l'apprentissage automatique permet de détecter les défaillances et les pannes potentielles avant qu'elles ne se produisent, ce qui permet d'optimiser la maintenance et de réduire les coûts.

La nouvelle approche de pronostic basée sur le modèle cloud computing et le principe de multitenancy présentée ici offre une solution de pronostic efficace à la demande d'un client, tout en assurant une meilleure qualité de service. En utilisant des méthodes de pronostic guidé par les données telles que les réseaux de neurones artificiels, les systèmes neuro-flous et les réseaux bayésiens, cette approche permet de sélectionner la méthode la plus efficace pour le client.

Cette approche permet également de mettre à jour régulièrement les modèles de pronostic en fonction des nouvelles données, ce qui permet d'améliorer la précision des prévisions. De plus, le fait de proposer le pronostic en tant que service permet de fournir une solution personnalisée et adaptée aux besoins spécifiques du client, sans avoir à investir dans l'infrastructure informatique nécessaire.

En somme, cette nouvelle approche de pronostic basée sur le modèle cloud computing et le principe de multitenancy offre une solution efficace et personnalisée de maintenance prédictive, ce qui peut aider à améliorer l'efficacité et la rentabilité des programmes de maintenance industrielle.[49]

2.10 Mesure de performances du Machine Learning

Une Confusion Matrix (matrice de confusion) ou tableau de contingence est un outil permettant de mesurer les performances d'un modèle de Machine Learning en vérifiant

notamment à quelle fréquence ses prédictions sont exactes par rapport à la réalité dans des problèmes de classification.

Le Machine Learning consiste à nourrir un algorithme à l'aide de données pour permettre à un ordinateur d'apprendre à effectuer une tâche spécifique. Afin de mesurer les performances d'un modèle de Machine Learning, on utilise généralement la Confusion Matrix ou matrice de confusion, aussi appelée tableau de contingence.

Outre le Machine Learning, les matrices de confusion sont aussi utilisées dans le domaine des statistiques, du Data Mining et de l'intelligence artificielle. De manière générale, elles permettent d'analyser des données statistiques plus rapidement et de rendre les résultats plus simples à déchiffrer via la Data Visualisation. Elles offrent l'opportunité d'analyser les erreurs dans les statistiques, le forage de données, ou même les examens médicaux.

2.10.1 Comment calculer une Confusion Matrix ?

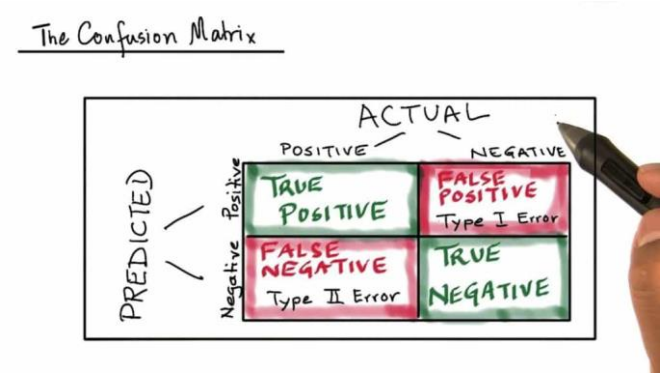


Figure 2-17 Matrice de confusion.[50]

Pour calculer une matrice de confusion, il est nécessaire de **disposer d'un ensemble de données de test (test dataset)** ou d'un ensemble de données de validation (validation dataset) avec les valeurs de résultat attendues. On fait ensuite une prédiction pour chaque ligne du » test dataset «.

A partir des résultats escomptés et des prédictions, la matrice indique le nombre de prédictions correctes pour chaque classe et le nombre de prédictions incorrectes pour chaque classe organisées en fonction de la classe prédite. Chaque ligne du tableau correspond à une classe prédite, et chaque colonne correspond à une classe réelle.

Dans les lignes sous les classes réelles, les prédictions ou les résultats sont inscrits. Ces résultats peuvent être l'indication correcte d'une prédiction positive comme » vraie positive » (true positive) et d'une prédiction négative comme » vraie négative » (true negative), ou une prédiction positive incorrecte comme » fausse positive » (false positive) et une prédiction négative incorrecte comme » fausse négative » (false negative).

L'avantage de ces matrices est qu'elles sont très simples à lire et à comprendre. Elles permettent de visualiser très rapidement les données et les statistiques afin d'analyser les performances d'un modèle et d'identifier les tendances qui peuvent aider à modifier les paramètres. Pour **les**

problèmes de classification avec trois classes ou plus, il est aussi possible d'utiliser une Confusion Matrix en ajoutant des lignes et des colonnes.

2.10.2 Comprendre les terminologies de la Confusion Matrix

Pour bien comprendre le fonctionnement d'une matrice de confusion, il convient de **bien comprendre les quatre terminologies principales** : TP, TN, FP et FN. Voici la définition précise de chacun de ces termes :

- **TP (True Positives)** : les cas où la prédiction est positive, et où la valeur réelle est effectivement positive. Exemple : le médecin vous annonce que vous êtes enceinte, et vous êtes bel et bien enceinte.
- **TN (True Negatives)** : les cas où la prédiction est négative, et où la valeur réelle est effectivement négative. Exemple : le médecin vous annonce que vous n'êtes pas enceinte, et vous n'êtes effectivement pas enceinte.
- **FP (False Positive)** : les cas où la prédiction est positive, mais où la valeur réelle est négative. Exemple : le médecin vous annonce que vous êtes enceinte, mais vous n'êtes pas enceinte.
- **FN (False Negative)** : les cas où la prédiction est négative, mais où la valeur réelle est positive. Exemple : le médecin vous annonce que vous n'êtes pas enceinte, mais vous êtes enceinte.

Recall (rappel)

$$Recall = \frac{TP}{TP + FN}$$

L'équation ci-dessus peut être expliquée en disant, à partir de toutes les classes positives, combien nous avons prédit correctement.

Le rappel doit être le plus élevé possible.

Precision

$$Precision = \frac{TP}{TP + FP}$$

L'équation ci-dessus peut être expliquée en disant, parmi toutes les classes que nous avons prédites comme positives, combien sont réellement positives.

La précision doit être aussi élevée que possible.

Précision

De toutes les classes (positives et négatives), combien d'entre elles avons-nous prédites correctement.

La précision doit être aussi élevée que possible.

F-measure

$$F - measure = \frac{2*Recall*Precision}{Recall + Precision}$$

Il est difficile de comparer deux modèles avec une faible précision et un rappel élevé ou vice versa. Donc, pour les rendre comparables, nous utilisons F-Score. Le score F aide à mesurer le rappel et la précision en même temps. Il utilise la moyenne harmonique à la place de la moyenne arithmétique en punissant davantage les valeurs extrêmes.[50]

2.11 Conclusion

Dans ce chapitre nous avons présenté dans la deuxième partie une définition et une vue générale sur machine learning et les algorithmes. Avant de commencer à créer une application d'apprentissage automatique, le choix d'une technologie parmi les nombreuses options disponibles peut s'avérer une tâche difficile. Par conséquent, il est important d'évaluer plusieurs options avant de prendre une décision finale. En outre, l'apprentissage du fonctionnement des différentes technologies d'apprentissage automatique.

Chapitre 3 : Etude de Cas

3.1. Introduction

L'exploration de données est en effet une étape cruciale du processus de découverte des connaissances. Elle consiste à analyser les données brutes en vue de trouver des motifs et des relations significatifs entre les variables et nous exposons les résultats de l'implémentation de 2 algorithmes d'apprentissage (régression linéaire, Random Forest) et une comparaison. Dans la première section, on décrit l'environnement de développement utilisé, et nous détaillons ses différentes étapes, les algorithmes choisis pour la modélisation de nos objectifs, ainsi que les différents outils permettant d'évaluer les performances des modèles produits. Le reste du chapitre est consacré à l'évaluation des capacités prédictives des modèles sur les jeux de données et à la comparaison de leurs résultats en utilisant les métriques appropriées. Enfin, la performance obtenue est évaluée à l'aide des modèles sélectionnés.

3.2. Environnement de développement

3.2.1. Environnement matériel

Nos expériences ont été réalisées sur un ordinateur portable dont les caractéristiques de l'environnement matériel sont rapportées dans le tableau 3.1

	Caractéristiques
Processeur	Intel(R) Core (TM) i7-9750H CPU @ 2.60GHz 2.59 GHz
RAM	16.0 GB
Système	64-bit operating system, x64-based processor
SSD	1 TB

Tableau 3.1 Caractéristiques de l'environnement matériel utilisé.

3.2.2. Environnement logiciel

L'utilisation d'environnements logiciels spécialisés dans la science des données et l'apprentissage automatique est devenue courante ces dernières années. Dont l'objectif commun est de faciliter le processus complexe d'analyse des données et de proposer des environnements intégrés en plus des langages de programmation standard. Les différents outils logiciels utilisés durant ce projet sont fournis par Google Colab.

Google Colab

Google Colab est une plateforme de notebooks Jupyter hébergée par Google. Elle offre un environnement de développement en ligne permettant d'exécuter du code Python, de créer et de partager des documents interactifs. Voici quelques points importants à savoir sur Google Colab :

- **Accès gratuit :** Google Colab est gratuit d'utilisation. Il fournit un accès aux ressources de calcul, y compris les processeurs et les GPU, sans frais supplémentaires. Toutefois, certaines fonctionnalités avancées peuvent nécessiter un abonnement payant à Google Cloud Platform.

- **Environnement de notebook interactif :** Google Colab permet de créer et d'exécuter des notebooks Jupyter, qui combinent du code Python, des textes explicatifs, des visualisations et des résultats. Cela facilite le partage et la collaboration sur des projets de programmation et d'analyse de données.
- **Accès aux GPU :** Colab offre la possibilité d'utiliser des GPU, notamment les GPU NVIDIA Tesla K80, T4 et P100. Cela permet d'accélérer l'exécution de tâches intensives en calcul, telles que l'apprentissage profond et l'entraînement de modèles d'apprentissage automatique.
- **Stockage et intégration aux services Google :** Colab est intégré aux services de stockage de Google, notamment Google Drive. Cela permet de sauvegarder et de charger facilement des données, des notebooks et des modèles depuis et vers le cloud.
- **Installation de dépendances :** Colab offre un environnement préinstallé avec de nombreuses bibliothèques populaires telles que NumPy, Pandas et TensorFlow. Cependant, si vous avez besoin de bibliothèques supplémentaires, vous pouvez les installer directement à partir du notebook à l'aide de commandes telles que : pip install.
- **Limite de temps d'exécution :** Les sessions sur Colab ont une limite de temps d'exécution, généralement de 12 heures. Après cette période, la session est réinitialisée et toutes les données non sauvegardées sont perdues. Cependant, le code et les résultats peuvent être sauvegardés dans des notebooks pour une utilisation ultérieure.

Google Colab est populaire parmi les chercheurs, les étudiants et les développeurs en raison de sa facilité d'utilisation, de son accès gratuit aux ressources de calcul et de sa capacité à exécuter du code Python dans un environnement interactif.[51]

3.2.3. Pytorch

PyTorch est un logiciel
Type : Bibliothèque logicielle
Développeur : Ronan Collobert puis équipes de recherche de Facebook.
Environnement(s) : Linux, Microsoft Windows et MacOS.
Langage de développement : C++, Python, C et Compute United Device Architecture.

Python : Python est également une plateforme open source, est devenu le langage de programmation le plus utilisé dans le domaine de la science des données et de l'apprentissage automatique. Sa popularité est due en partie à sa syntaxe simple et intuitive, qui permet aux programmeurs d'écrire du code clair et facile à comprendre. Python est également un langage de haut niveau, ce qui signifie qu'il est plus abstrait que les langages de bas niveau comme le langage C, ce qui facilite la programmation. Enfin, la version de Python utilisée dans ce travail est la version 3.9, qui est une version stable et largement utilisée de Python. Python 3.9 a été publié en Oct.5,2020 et inclut de nombreuses fonctionnalités nouvelles et améliorées par rapport aux versions précédentes de Python.

Python est un langage de programmation polyvalent et populaire qui est utilisé dans de nombreuses applications, telles que :

- ✓ La programmation d'applications.
- ✓ La création de services web.
- ✓ La génération de code.
- ✓ La métaprogrammation.[52]

3.3. Description des données utilisées

Les données recueillies sur le colmatage des filtres à particules et les mesures de débit et de pression différentielle à travers le filtre sont des informations précieuses pour évaluer l'état de fonctionnement du filtre et prédire les besoins futurs en matière de maintenance. L'utilisation de capteurs pour enregistrer le débit et la pression différentielle permet de surveiller en temps réel les performances du filtre. Le débit correspond au volume de gaz qui passe à travers le filtre, tandis que la pression différentielle mesure la différence de pression entre l'entrée et la sortie du filtre. En analysant les données recueillies au fil du temps, il est possible de détecter les signes de colmatage du filtre. Le colmatage se produit lorsque les particules solides s'accumulent dans le filtre, réduisant ainsi le débit d'air et augmentant la pression différentielle. En surveillant ces paramètres, on peut déterminer quand le filtre atteint un niveau critique de colmatage, nécessitant ainsi une intervention de maintenance.

Le passage d'une stratégie de maintenance préventive à une maintenance prédictive est rendu possible grâce à l'analyse des données collectées. En utilisant des techniques d'apprentissage automatique et de modélisation prédictive, il est possible de développer des algorithmes capables de prédire le moment optimal pour effectuer la maintenance sur le filtre. Ces prédictions peuvent être basées sur des modèles qui établissent des relations entre le colmatage du filtre, le débit, la pression différentielle et d'autres facteurs influençant les performances du filtre.

3.3.1. Les défis spécifiques (data-set)

Les data-sets (ensembles de données) sont des collections structurées de données utilisées pour l'entraînement, la validation et les tests dans le domaine de l'apprentissage automatique (machine Learning) et de l'intelligence artificielle. Posent généralement plusieurs défis spécifiques qui nécessitent d'être traités avant d'être utilisés dans la modélisation. Nous penchons sur ces problèmes et proposons des méthodes pour les résoudre.

- Train Data (Données d'entraînement) : Ce fichier contient les données des capteurs enregistrent le débit et la différence de pression à travers le filtre.
- Test Data (Données de test) : Ce fichier contient les données des capteurs qui n'ont pas enregistré d'événements de défaillance. Ces données seront utilisées pour tester les performances des modèles prédictifs formés à l'aide des données d'entraînement.

L'objectif sera donc d'utiliser les données d'entraînement pour développer des modèles qui peuvent prédire la durée de vie utile restante des filtres avant leur défaillance. Ces modèles seront ensuite évalués en utilisant les données de test. Ce processus permettra de mesurer la performance des modèles de maintenance prédictive et de déterminer leur capacité à prédire avec précision les défaillances des filtres.

Id : est l'identificateur (numéro) des filtres 1 à 50.

Pression différentielle : La différence de pression entre deux points d'un système (filtre)

Débit : il fait référence à la quantité de fluide, de gaz ou de particules qui passe à travers un système dans une unité de temps spécifiée.

Temps : et il représente la durée d'un événement ou d'une action.

Alimentation en poussière : débit de poussière, qui représente la quantité de poussière qui passe à travers les filtres.

Les données sont disponibles en accès libre sur le site Kaggle à l'adresse suivante

https://www.kaggle.com/datasets/prognosticshse/preventive-to-predicitve-maintenance

La table ci-dessous illustre un sous ensemble des données d'entraînement dont les colonnes représentent les paramètres suivants :

Id	Pd	Débit	Temps	Ap
1	**0**	**0**	**0.1**	**236.429**
1	**0**	**0**	**0.2**	**236.429**
1	**0**	**0**	**0.3**	**236.429**
1	**0**	**0**	**0.4**	**236.429**
1	**0**	**0**	**0.5**	**236.429**
50	**359.972**	**58.7219**	**59.4**	**177.322**
50	**360.786**	**58.6999**	**59.5**	**177.322**
50	**361.509**	**58.7438**	**59.6**	**177.322**
50	**362.052**	**58.6012**	**59.7**	**177.322**
50	**366.482**	**58.6121**	**59.8**	**177.322**

Tableau 3-2 Exemple de données d'entraînement.

3.3.2. Prétraitement des données

Tout d'abord, nous détaillons la méthode utilisée pour créer les ensembles de données essentiels à l'entraînement des modèles de prédiction, ainsi que l'ensemble de données utilisé pour évaluer la qualité des prédictions générées par ces modèles. Ensuite, nous abordons le processus de sélection des paramètres optimaux pour nos modèles et expliquons comment cette sélection a été effectuée.

3.3.3. Entraînement et paramétrage des modèles

Tout d'abord, nous expliquons comment générer l'ensemble de données nécessaire pour l'entraînement des modèles prédictifs et l'ensemble de données à utiliser pour l'évaluation de la qualité des prédictions générées par ces modèles. Ensuite, nous discutons la façon dont nous avons choisi les paramètres optimaux pour nos modèles.

- **Sélection des données d'entraînement et de test**

On divise les données en deux ensembles :

- Ensemble d'entraînement (Training data) : Il représente les (39420) premiers cas de l'ensemble des données qui seront utilisées pour entraîner les modèles.
- Ensemble de test (Test data) : Il est constitué des 50 cas des données restants, destiné à l'évaluation de la performance prédictive des modèles.

- **Paramétrage des modèles**

Afin d'obtenir des performances élevées d'un modèle, il est nécessaire de choisir les meilleurs paramètres pour l'algorithme qui génère ce modèle. Ce processus peut être réalisé à l'aide de deux méthodes. Dans le cas de la régression, ces paramètres sont généralement sélectionnés manuellement, tandis que dans la classification, ils sont calculés automatiquement en utilisant une méthode appelée "grid search". Le grid search effectue une recherche exhaustive dans un sous-ensemble spécifié manuellement de l'espace des hyper paramètres de l'algorithme ciblé. Une fois l'apprentissage terminé, les performances du modèle sont évaluées et les paramètres qui offrent les meilleures performances sont identifiés. Ces paramètres optimaux doivent ensuite être utilisés pour toutes les données d'entraînement et de test.

3.4. Approche proposée

Dans cette section, nous citons brièvement les algorithmes proposés (Random Forest et régression linéaire) pour modéliser certains objectifs de la maintenance prédictive sur l'ensemble de données préparé plus haut (voir figure 3-1)

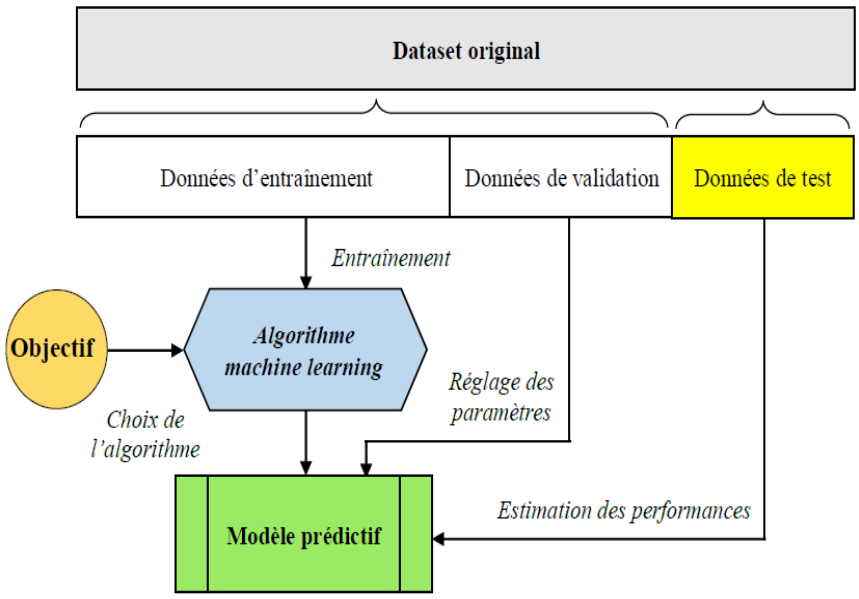


Figure 3-1 Les étapes de création d'un modèle prédictif.

3.4.1. Objectifs

Les objectifs de la maintenance prédictive traités dans ce projet sont :

1- Régression : Prédire le temps de fonctionnement avant panne (TTF) d'un filtre.

2- Déterminer si un filtre tombera en panne pendant une période ou non.
3- Identifier dans quelle période un filtre tombera en panne.

3.5. Choix des algorithmes d'apprentissage

Dans ce chapitre, nous présentons plusieurs algorithmes classiques adaptés aux problèmes de classification et de régression. Ces algorithmes, dont les principes fondamentaux ont été abordés dans le chapitre précédent, répondent à nos objectifs principaux. Le chapitre suivant se concentrera sur l'implémentation de ces algorithmes et l'évaluation de leurs performances à travers les résultats obtenus.

3.5.1. Modèles de régression

Les modèles de régression en maintenance prédictive sont employés pour estimer la période de temps restante avant qu'un actif ne nécessite une réparation ou ne tombe en panne. Cela permet de prédire la durée de vie utile restante de l'actif, c'est-à-dire la période pendant laquelle l'actif peut continuer à fonctionner sans problème avant de présenter une défaillance Algorithmes : Régression linéaire, Random Forest ,KNN, SVM.

	Id	Dp	débit	temps	Ap
count	**39420.000**	**39420.000**	**39420.000**	**39420.0000**	**39420.0000**
mean	**25.167453**	**78.232050**	**66.095780**	**49.135490**	**132.247854**
std	**14.315149**	**107.342894**	**11.306582**	**39.088269**	**62.716754**
Min	**1.000000**	**0.000000**	**0.000000**	**0.100000**	**59.107236**
25%	**12.000000**	**8.590133**	**58.458498**	**19.800000**	**79.246266**
50%	**24.000000**	**35.174340**	**59.040107**	**39.900000**	**118.214472**
75%	**38.000000**	**102.358200**	**81.136739**	**65.800000**	**158.492533**
Max	**50.000000**	**607.910200**	**84.199355**	**179.400000**	**316.985065**

Tableau 3-2 Données d'entraînement.

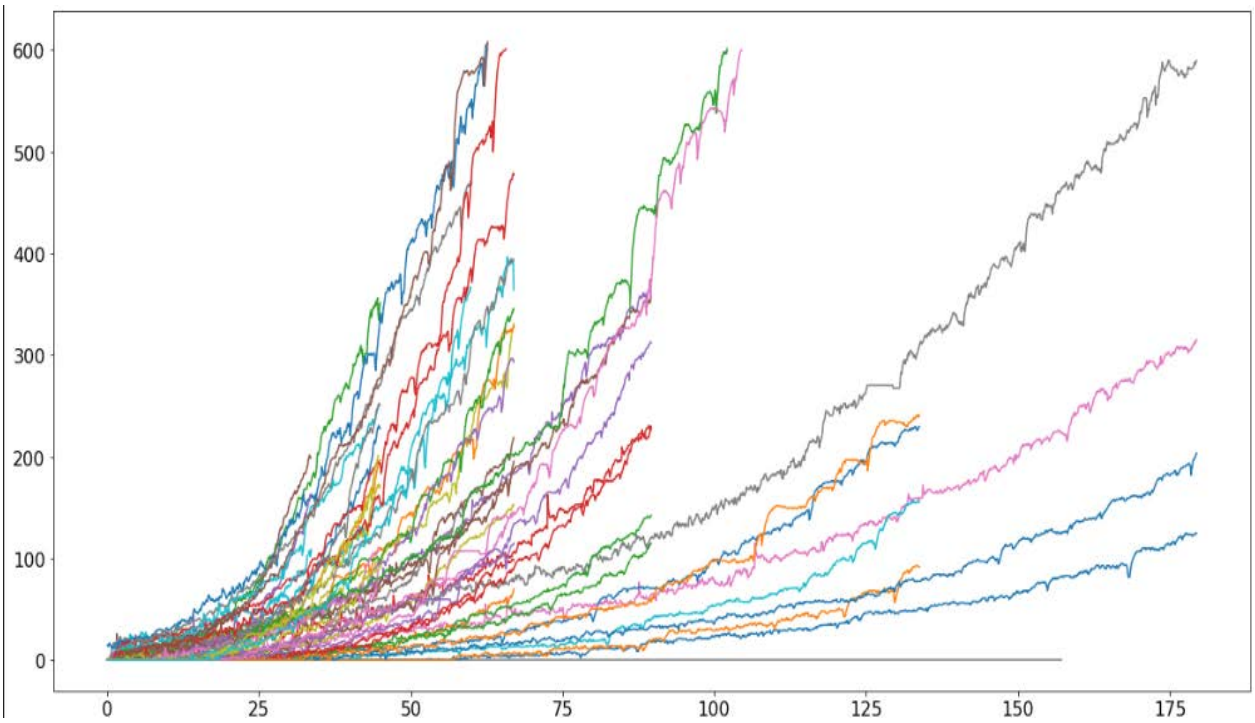


Figure 3-2 Pressions différentielles dans le temps.

Nous pouvons voir que certaines séries chronologiques atteignent 600. Il s'agit des données d'échec de la rame. Sur ces courbes, nous voulons prédire, lorsqu'il atteint 600 = échec.

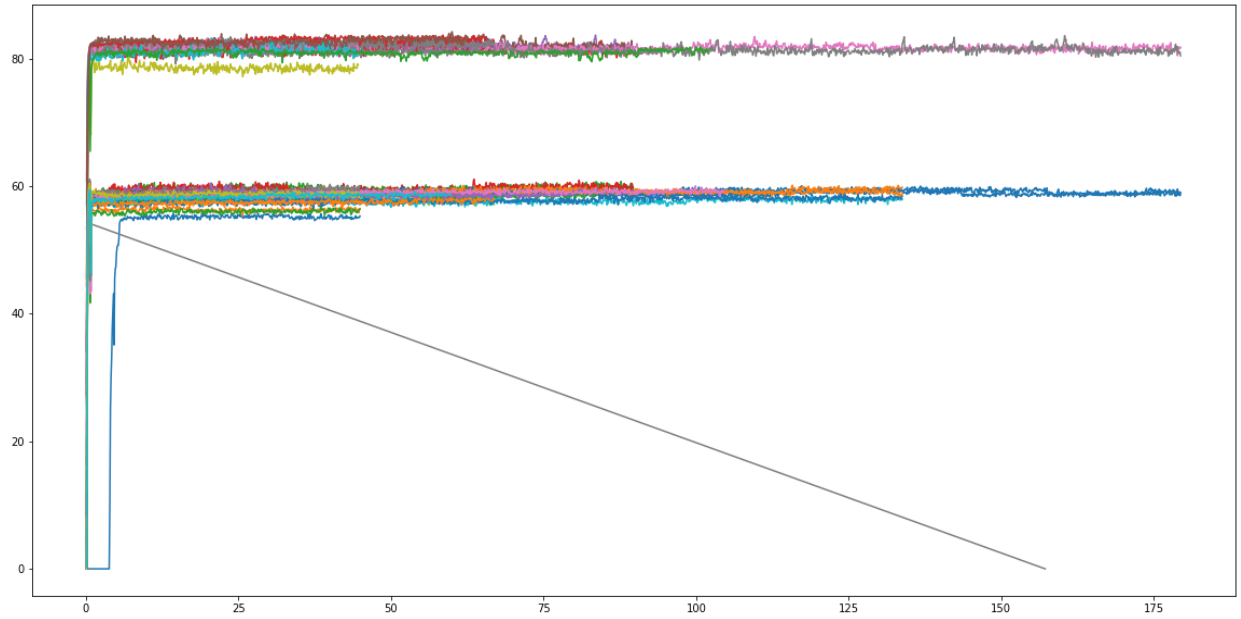


Figure 3-3 La différence entre deux débits.

Il existe deux débits différents mais cela ne change pas dans le temps. Le débit est constant.

	Pression différentielle	Débit	Alimentation en poussière	RUL
Pression différentielle	**1.000000**	**0.162771**	**-0.082623**	**-0.458465**
Débit	**0.162771**	**1.000000**	**-0.411261**	**0.273620**
Alimentation en poussière	**-0.082623**	**-0.411261**	**1.000000**	**-0.664177**
RUL	**-0.458465**	**0.273620**	**-0.664177**	**1.000000**

Tableau 3-4 La relation entre Variables.

La corrélation des données avec la matrice de corrélation :

La matrice de corrélation est un tableau qui présente de manière systématique les corrélations entre les variables d'un ensemble de données. Elle est largement utilisée pour analyser les relations linéaires entre les variables et détecter les schémas ou les dépendances qui peuvent exister entre elles. L'analyse de la matrice de corrélation permet d'obtenir des informations importantes sur la structure des données et les relations entre les variables.

1. **Valeurs de corrélation :** Les valeurs de corrélation varient généralement entre -1 et 1. Une corrélation de 1 indique une corrélation positive parfaite, ce qui signifie que les variables augmentent ou diminuent ensemble de manière linéaire. Une corrélation de -1 indique une corrélation négative parfaite, ce qui signifie que les variables ont une relation linéaire inverse. Une corrélation proche de zéro indique une faible corrélation linéaire entre les variables.
2. **Force de corrélation :** La magnitude de la corrélation indique la force de la relation entre les variables. Une corrélation proche de 1 ou -1 indique une forte relation linéaire, tandis qu'une corrélation proche de 0 indique une faible relation linéaire.
3. **Direction de la corrélation :** Le signe de la corrélation indique la direction de la relation entre les variables. Une corrélation positive signifie que les variables augmentent ou diminuent ensemble, tandis qu'une corrélation négative signifie qu'elles ont une relation inverse.
4. **Patterns de corrélation :** L'analyse de la matrice de corrélation peut révéler des schémas ou des structures dans les données. Par exemple, des regroupements de variables fortement corrélées peuvent indiquer des relations thématiques ou des dépendances mutuelles. Des corrélations élevées entre une variable spécifique et plusieurs autres variables peuvent indiquer une influence importante de cette variable sur les autres.
5. **Prudence avec la causalité :** Il est essentiel de noter que la corrélation ne signifie pas nécessairement causalité. Une forte corrélation entre deux variables ne signifie pas

automatiquement qu'une variable cause l'autre. Il est possible que des facteurs confondants ou des relations indirectes influencent la corrélation observée. Il est important de mener des études supplémentaires pour établir des relations de causalité entre les variables.

3.6. Algorithmes d'apprentissage automatique

3.6.1. La régression

Dans cette partie, des modèles de régression linéaire et non linéaire ont été créés pour prédire la durée de vie restante d'un filtre. Les algorithmes d'apprentissage automatique proposés ont été essayées et leurs mesures de performance ont été calculées et évaluées. Nous avons estimé le RUL pour 50 filtres. Pour valider nos modèles, nous avons comparé les performances des métriques de la régression des différents modèles en utilisant l'ensemble de caractéristiques originales. Les principales mesures d'évaluation de la régression calculées pour chaque modèle étaient chaque modèle étaient l'erreur quadratique moyenne (RMSE), le R au carré (R2), l'erreur absolue moyenne (MAE) et EVS expliquée.

3.6.2. Métriques de régression

Pour évaluer les modèles de régression et de les comparer, on peut calculer la distance entre valeurs prédites et vraies valeurs. Cela nous donne plusieurs critères :

- **L'erreur quadratique moyenne (Root mean squared error)**

L'erreur quadratique moyenne (RMSE) est une mesure couramment utilisée pour évaluer les performances d'un modèle de régression. Elle représente la racine carrée de la moyenne des erreurs quadratiques entre les valeurs prédites par le modèle et les valeurs réelles.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(p_i - a_i)^2}{n}}$$

a = Cible réelle.

p = Cible prévue.

- **L'erreur absolue moyenne (Mean Absolute error)**

L'erreur absolue moyenne (MAE) à la même unité que les données d'origine et ne peut être comparé qu'entre des modèles dont les erreurs sont mesurées dans les mêmes unités. Son ampleur est généralement similaire à celle du RMSE, mais légèrement plus petite. a et p sont défini dans l'erreur quadratique moyenne.

$$MAE = \frac{\sum_{i=1}^{n} | p_i - a_i |}{n}$$

➢ Le coefficient de détermination R^2

Le coefficient de détermination R2, également connu sous le nom d'erreur quadratique moyenne (R-squared error), mesure le pouvoir explicatif d'un modèle de régression. Il indique la proportion de la variance de la variable dépendante (y) qui peut être expliquée par les variables indépendantes (x) incluses dans le modèle.

Le coefficient de détermination R2 est calculé à partir des termes des sommes des carrés, tels que :

$$R^2 = 1 - \frac{SSE}{SST}$$

Où SSE représente la somme des carrés des résidus (erreur de prédiction) et SST représente la somme des carrés totale.

Le coefficient de détermination R2 est une valeur comprise entre 0 et 1. Une valeur de 0 indique que le modèle n'explique aucune variation de la variable dépendante, tandis qu'une valeur de 1 indique que le modèle explique parfaitement toute la variation. Une valeur proche de 1 indique une bonne adéquation du modèle aux données observées.

➢ EVS (Explained Variance Score)

L'EVS est une métrique qui évalue la capacité d'un modèle à expliquer la variance des données observées. Il est calculé en comparant la variance des prédictions du modèle à la variance des données réelles. Une valeur de 1 indique que le modèle explique parfaitement la variance des données, tandis qu'une valeur de 0 indique que le modèle n'explique pas du tout la variance.

3.7. Résultats obtenus

Les résultats finaux des métriques sont présentés dans les tableaux suivants.

3.7.1. Régression linéaire

	Train	Test
MAE	**3.670730039968789**	**4.65440614047318**
R2	**0.9965893402205237**	**0.99961945249884**
RSME	**6.0279494662841735**	**7.472944577450328**
EVS	**0.996589340220537**	**0.9996229019681166**

Tableau 3-5 Evaluation de la régression linéaire.

3.7.2. Support Vector

	Train	Test
MAE	**4.085472516315362**	**4.994717745017307**
R2	**0.9929890451776926**	**0.9946425135054054**

Tableau 3-6 Evaluation du Support Vector.

3.7.3. Random forest

	train	test
MAE	**1.0062570609570065**	**4.54995527151493**
R2	**0.9997274141381588**	**0.9961945465202039**
RMSE	**1.704127427187089**	**7.42923436041596**
EVS	**0.9997276030758042**	**0.9962801722058664**

Tableau 3-7 Evaluation de Random forest.

3.7.4. KNN

	Train	Test
MAE	**2.059199818807296**	**4.998602222645428**
R2	**0.9988679384262158**	**0.9949747238115659**

Tableau 3-8 Evaluation de KNN.

Évaluation du meilleur modèle

Le meilleur modèle est le Random Forest Regressor.

3.7.5.Comparaison :

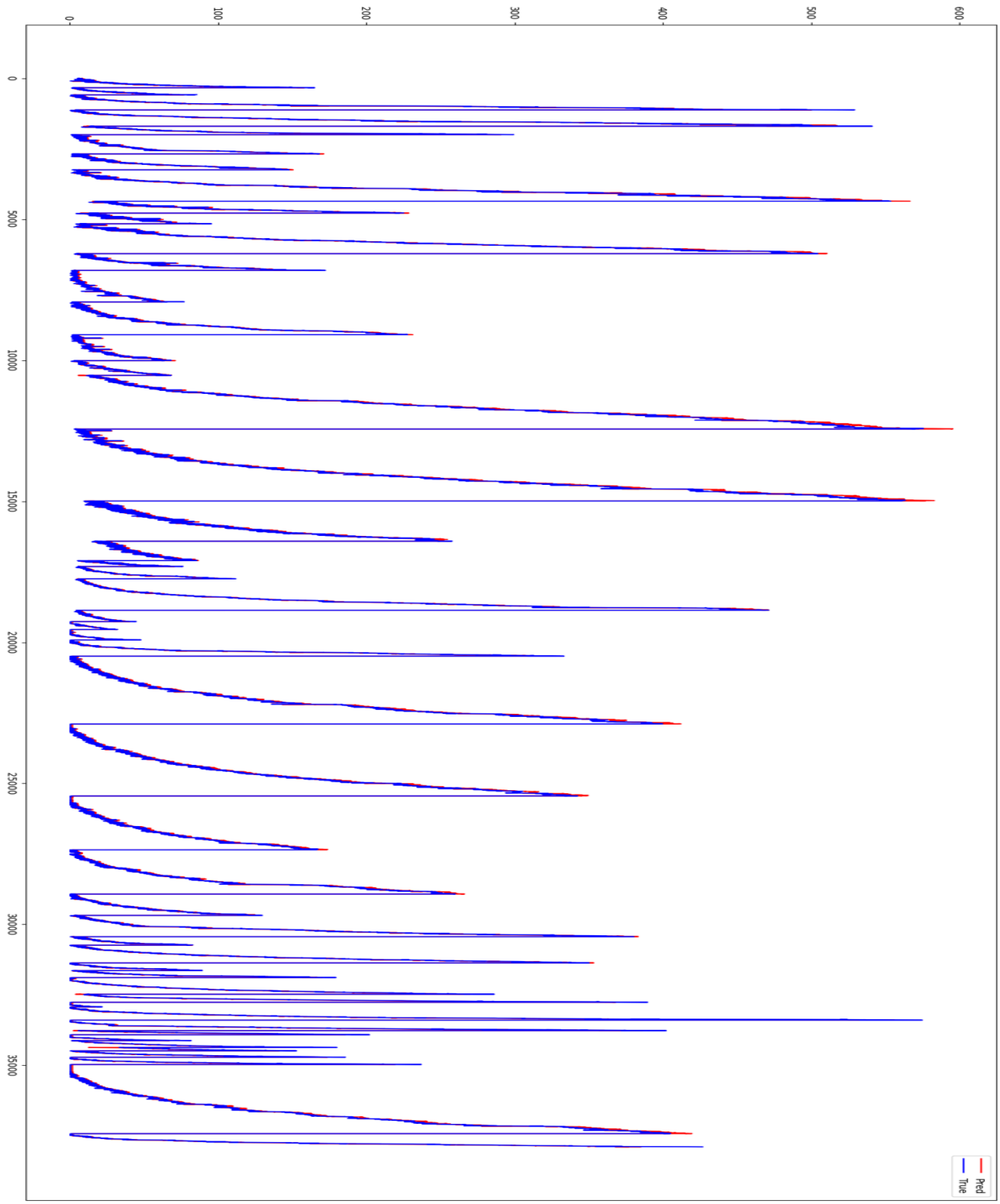


Figure 3-4 Comparaison entre la valeur réelle et prédictive.

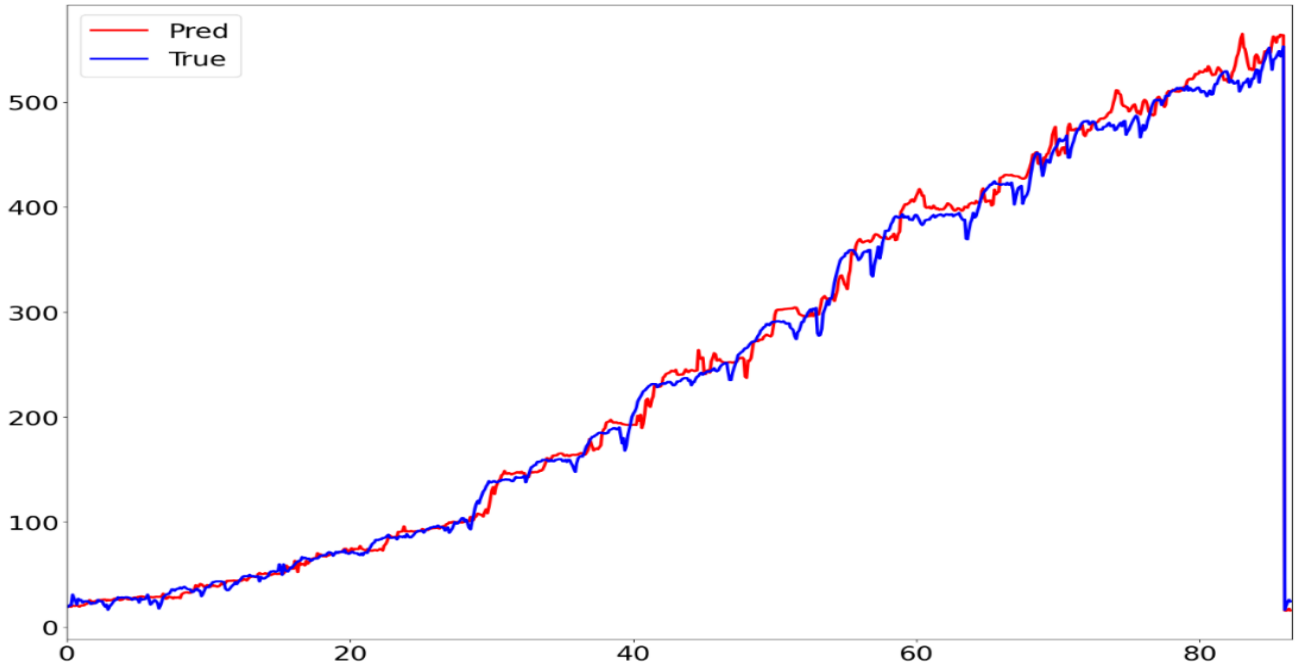


Figure 3-5 Pression différentielle de la valeur réelle et prédite

3.8. Conclusion

Dans ce chapitre, nous avons présenté le jeu de données qui utilisé pour développer, tester et évaluer une approche de maintenance prédictive. Ensuite, nous avons décrit les opérations de prétraitement effectuées sur le jeu de données afin de le préparer pour la phase d'apprentissage. Enfin, nous avons abordé brièvement une approche d'apprentissage automatique basée sur la régression et la classification, ainsi que les différentes métriques pouvant être utilisées pour évaluer le modèle d'apprentissage. Nous avons également procédé à l'implémentation de quelques algorithmes d'apprentissage et évalué leurs performances.

Conclusion Générale

Conclusion Générale

Cette étude se concentre sur l'application de l'intelligence artificielle à la maintenance prédictive, plus précisément dans le contexte des filtres à air pour véhicules roulants. L'objectif est de prédire la durée de vie des filtres, d'identifier les filtres susceptibles d'être défectueux pendant une période donnée et de prédire les périodes auxquelles les pannes se produiront, afin de planifier les opérations de maintenance. Pour cela, nous avons réalisé une analyse approfondie de la maintenance dans un premier temps, ce qui nous a permis de comprendre les différentes techniques de maintenance disponibles ainsi que leurs avantages et inconvénients. Ensuite, nous avons effectué une revue de l'état de l'art dans le domaine de l'apprentissage automatique (Machine Learning) pour identifier les algorithmes appropriés pour modéliser notre problématique. Nous avons également recherché les outils qui nous permettront de tester l'efficacité des solutions obtenues. En combinant les connaissances acquises en matière de maintenance et les techniques d'apprentissage automatique, nous visons à développer des modèles prédictifs qui aideront à optimiser la gestion des filtres à air, en anticipant les pannes et en planifiant les opérations de maintenance de manière plus efficace.

Le projet présenté dans ce mémoire propose une solution visant à exploiter le potentiel des modèles d'intelligence artificielle pour améliorer le processus de maintenance industrielle en prévoyant les défauts des machines à l'avance. Nous avons souligné l'importance d'avoir un programme de maintenance prédictive efficace. En effet, les arrêts imprévus des chaînes de production et les retards de production entraînent des pertes financières et temporelles qui affectent directement la santé budgétaire et la compétitivité des entreprises. Cependant, pour obtenir de bons résultats ou des prédictions fiables, il est crucial de disposer d'une vaste base de données contenant de nombreuses données brutes collectées sur les équipements. En effet, plus un modèle dispose de données d'entraînement, plus il est en mesure de prédire correctement les observations. De plus, ces données doivent être cohérentes et de haute qualité. Il est à noter que la prédiction à partir de jeux de données simulés est largement réalisable de nos jours grâce au développement et au partage en open source de plates-formes de données. De plus, le développement de nouvelles bibliothèques Python telles que Scikit-learn a considérablement facilité l'application de l'intelligence artificielle dans le domaine de la maintenance, en particulier pour les personnes non spécialistes dans ce domaine.

En conclusion, l'utilisation de l'apprentissage automatique offre aux responsables de la maintenance des informations précieuses pour comprendre, améliorer et préparer les ressources matérielles et humaines nécessaires avant qu'une défaillance ne survienne. Cela permet de remplacer les stratégies de maintenance traditionnelles, telles que la maintenance corrective et préventive, par une approche de maintenance prédictive basée sur des modèles prédictifs. L'évaluation des performances des modèles obtenus démontre que l'application de l'intelligence artificielle dans le domaine industriel est extrêmement bénéfique pour les entreprises. Elle permet d'augmenter la rentabilité de leurs équipements de production et, par conséquent, d'accroître les bénéfices de l'entreprise. En utilisant des modèles prédictifs, les entreprises peuvent anticiper les pannes, planifier efficacement les opérations de maintenance et éviter les arrêts non planifiés de la production. Cela permet d'optimiser l'utilisation des ressources, de minimiser les pertes financières

et de renforcer la compétitivité sur le marché. L'intelligence artificielle offre donc une nouvelle approche puissante pour améliorer la gestion de la maintenance industrielle et maximiser les performances opérationnelles.

Lors de la conception et de la réalisation de notre projet, nous avons rencontré plusieurs défis que nous avons réussi à surmonter, ce qui a été extrêmement bénéfique. Tout d'abord, nous avons dû acquérir de nouvelles connaissances dans des domaines tels que la science des données et l'apprentissage automatique. Ensuite, lors de l'implémentation de notre solution, nous avons découvert nouveaux outils de programmation tels que Python, Jupyter Notebook, PyCharm, ainsi que différentes bibliothèques permettant de traiter les données, de mettre en œuvre des algorithmes d'apprentissage automatique et de visualiser les résultats (Numpy, Pandas, Matplotlib, Scikit-learn, etc.). Cette expérience nous a également permis d'améliorer nos compétences en méthodologie de recherche, en communication et en rédaction. Sur le plan personnel, ce projet ouvre de nombreuses perspectives d'amélioration. Nous avons travaillé sur des données simulées plutôt que sur des données réelles provenant de machines en utilisation réelle. Il est donc essentiel de procéder à une phase d'acquisition et de collecte de données dans un environnement industriel avant de mettre en œuvre le système.

Étant donné les limites et les limitations des techniques d'apprentissage automatique dans certaines situations, nous proposons d'utiliser des techniques d'apprentissage plus avancées, telles que l'apprentissage en profondeur (Deep Learning), pour modéliser plus efficacement les tâches de maintenance. Enfin, nous prévoyons de poursuivre nos travaux sur l'application de l'apprentissage automatique à la maintenance industrielle dans le cadre d'un doctorat, afin d'approfondir nos connaissances dans ce domaine fascinant et de contribuer aux avancées scientifiques par le biais de la recherche

Références bibliographiques

[01] l'AFNOR en 1994(norme NFX 60- e.010).
[02] NF EN 13306 la norme **AFNOR** (Norme européenne, Éditée et diffusée par Association Française de Normalisation) **NF-X 60 000** Juin 2001.
[03] https://www.researchgate.net/figure/Les-differents-types-de-maintenance_fig1_281793037
[04] Selon la définition AFNOR (norme x60-010).
[05] Norme x60-319/nf en 13306 terminologies de la maintenance, 2001.
[06]Cours maintenance industrielle université mostefa ben boulaid -batna2
[07]https://www.tribofilm.fr/les-differents-types-de-maintenance/
[08] GESTION DE LA MAINTENANCE INDUSTRIELLLE« Cours pour Master 1, Mécatronique ».
[09] Extrait norme NF EN 13306 X 60-319, AFNOR X 60-000.
[10] https://batiadvisor.fr/maintenance-predictive/
[11]https://expertise.boschrexroth.fr/tout-comprendre-sur-la-maintenance-predictive/ 2/2/2023
[12] https://www.spectraltms.com/blog/quest-ce-que-la-maintenance-predictive.8/3/2023
[13]https://www.journaldunet.fr/web-tech/dictionnaire-de-l-iot/1489507-maintenance-predictive-definition-et-interet-dans-l-industrie/
[14]Maëlys De Santis, « Qu'est-ce que la maintenance prédictive », URL : https://www.appvizer.fr/magazine/operations/gmao/maintenance-predictive-definition, 2/2/2023
[15] Innocent Mateyaunga, « Prédictive Maintenance Using Machine Learning », la thèse de master, sous la direction de Hadj Abdelkader, Faculté de technologie de l'université de Tlemcen.
[16] https://www.ibm.com/fr-fr/cloud/learn/machine 11/03/2023
[17] https://www.oracle.com/dz/artificial-intelligence/what-is-ai/ 11/03/2023
[18] https://www.sicara.fr/parlons-data/machine-learning 14/03/2023
[19]https://aws.amazon.com/fr/solutions/implementations/predictive-maintenance-using-machine-learning/ 14/03/2023
[20]https://praedictia.com/page/lapprentissage-machine/lhistoire-de-lapprentissage-machine.html12/03/2023
[21] https://myriad-data.com/wp-content/uploads/2019/07/nlp.pdf7 18/03/2023
[22] file:///C:/Users/BENZ/Downloads/sustainability-12-08211-v3%20(2).pdf 20/03/2023
[23] https://ia-data-analytics.fr/machine-learning/usages/23/03/2023
[24]https://newsinitiative.withgoogle.com/fr-fr/resources/lessons/different-approaches-to-machine-learning/27/03/2023
[25] https://www.talend.com/fr/resources/what-is-machine-learning/ 1/04/2023
[26]https://www.journaldunet.fr/web-tech/guide-de-l-intelligence-artificielle/1501311-apprentissage-supervise/5/04/2023
[27] https://datascientest.com/apprentissage-non-supervise 10/04/2023
[28] https://brightcape.co/machine-learning-open-source/ 11/04/2023
[29] https://link.springer.com/article/10.1007/s10586-021-03240-4 18/04/2023
[30] https://www.actuia.com/keras/pourquoi-utiliser-keras/ 19/04/2023

Références bibliographiques

[31]https://www.quora.com/Which-deep-learning-framework-between-Tensorflow-or-PyTorch-should-I-choose-for-NLP-research 20/04/2023
[32]https://www.jedha.co/formation-ia/algorithme svm#:~:text=C'est%20le%20cas%20des,de%20discrimination%2C%20et%20de%20r%C3%A9gression.26/04/2023
[33] https://www.jedha.co/formation-ia/algorithmes-machine-learning 27/04/2023
[34]https://www.journaldunet.fr/web-tech/guide-de-l-intelligence-artificielle/1501879-machine-a-vecteurs-de-support-svm-definition-et-cas-d-usage/ 27/04/2023
[35]https://www.datacamp.com/tutorial/svm-classification-scikit-learn-python 27/04/2023
[36] https://aws.amazon.com/fr/what-is/linear-regression/
[37] https://www.analyticsvidhya.com/blog/2022/06/linear-regression-using-mlib/
[38] http://thatit.free.fr/COURS/Cours_7.pdf
[39] https://constellation.uqac.ca/id/eprint/5857/1/Zoungrana_uqac_0862N_10694.pdf
[40] https://haltode.fr/algo/ia/apprentissage_artificiel/regression_lin_poly.html
[41] https://serokell.io/blog/polynomial-regression-analysis
[42] https://www.ibm.com/fr-fr/topics/logistic-regression
[43]https://www.ibm.com/topics/randomforest#:~:text=Random%20forest%20is%20a%20commonly,both%20classification%20and%20regression%20problems.
[44]https://towardsdatascience.com/from-a-single-decision-tree-to-a-random-forest-b9523be65147
[45]https://www.researchgate.net/publication/319939107_Les_Reseaux_de_Neurones_Artificiels
[46]https://openclassrooms.com/fr/courses/5801891-initiez-vous-au-deep-learning/5814616-explorez-les-reseaux-de-neurones-en-couches
[47] https://datascience.eu/fr/apprentissage-automatique/perceptron/
[48] https://www.becoz.org/these/memoirehtml/ch06s04.html
[49] LES APPROCHES DE LA MAINTENANCE PREDICTIVE INTELLIGENTE DANS L'INDUSTRIE 4.0 UNIVERSITE BADJI MOKHTAR – ANNABAA
[50] https://www.lebigdata.fr/confusion-matrix-definition
[51]https://research.google.com/colaboratory/faq.html?hl=fr#:~:text=Colaboratory%2C%20souvent%20raccourci%20en%20%22Colab,donn%C3%A9es%20et%20%C3%A0%20l'%C3%A9ducation.
[52]https://www.journaldunet.fr/web-tech/dictionnaire-du-webmastering/1445304-python-definition-et-utilisation-de-ce-langage-informatique/

Résumé

La maintenance prédictive est en effet devenue une approche précieuse pour réduire les temps d'arrêt non planifiés et minimiser les coûts de maintenance élevés dans diverses industries. En utilisant des techniques avancées telles que la science des données et l'apprentissage automatique (machine learning), il est possible d'exploiter les énormes quantités de données générées par les filtres a air pour véhicules roulants simulés pour prédire leur état de fonctionnement futur. Le but de ce projet est d'exploiter une énorme quantité de données relatives au comportement des filtres à air pour véhicules roulants simulés afin d'entraîner des modèles capables de prédire l'état de fonctionnement futur de ces filtres. Ainsi, nous avons créé des modèles prédictifs pour estimer la durée de vie restante d'un filtre, trouver quels filtres tomberont en panne dans une période donnée, ainsi pour prédire la période pendant laquelle un filtre tombera en panne. Ces modèles sont générés par quatre algorithmes Régression Linéaire, Random Forest, SVM, KNN.

Mots clé : Maintenance prédictive, l'apprentissage automatique, Régression.

Abstract

Predictive maintenance has indeed become a valuable approach to reduce unplanned downtime and minimize high maintenance costs in various industries. By utilizing advanced techniques such as data science and machine learning, it is possible to leverage the enormous amount of data generated by simulated air filters for rolling vehicles to predict their future operational state. The goal of this project is to exploit a vast amount of data related to the behavior of simulated air filters for rolling vehicles in order to train models capable of predicting the future operational state of these filters. Thus, we have created predictive models to estimate the remaining lifespan of a filter, identify which filters will fail within a given period, and predict the timeframe in which a filter will fail. These models are generated using four algorithms: linear regression, random forest, SVM, KNN.

Keywords: Predictive maintenance, Machine learning, Regression.

ملخص

أصبحت الصيانة التنبؤية نهجًا قيّمًا لتقليل الأوقات غير المخططة للتوقف وتقليل التكاليف المرتفعة للصيانة في مختلف الصناعات. باستخدام تقنيات متقدمة مثل علم البيانات والتعلم الآلي (التعلم الآلي) ، من الممكن استغلال الكميات الهائلة من البيانات الناتجة عن محاكاة فلاتر هواء السيارات للتنبؤ بحالة التشغيل المستقبلية. الهدف من هذا المشروع هو استغلال كمية هائلة من البيانات المتعلقة بسلوك مرشحات الهواء للمركبات المتدحرجة المحاكاة من أجل تدريب نماذج قادرة على التنبؤ بحالة التشغيل المستقبلية. لذا ، قمنا بإنشاء نماذج تنبؤية لتقدير المدة الزمنية المتبقية للمرشح ، وللعثور على أي مرشحات ستتعطل في فترة زمنية معينة ، وللتنبؤ بالفترة التي ستتعطل فيها المرشح. تم إنشاء هذه النماذج باستخدام خوارزميتين ، الانحدار الخطي والغابة العشوائية، آلة الدعم بالفواصل والجار الأقرب K-plus.

الكلمات المفتاحية : الصيانة التنبيية، التعلم الآلي، الانحدار.