

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE MOHAMED BOUDIAF - M'SILA

FACULTE SCIENCES ET TECHNOLOGIE
DEPARTEMENT GENIE ELECTRIQUE



DOMAINE : SCIENCES ET TECHNOLOGIE
FILIERE : GENIE ELECTRIQUE
OPTION : COMMANDE ELECTRIQUE

Mémoire présenté pour l'obtention
Du diplôme de Master Académique

Par : MOUSSAI Saber & SALEM Chaima

Intitulé

Diagnostic Des Défauts D'une Turbine A Gaz Basé
Sur Des Techniques D'apprentissage Automatique

Soutenu devant le jury composé de :

Mr. Issam ZORIG	Université de M'Sila	Président
Mr. Abdelghafour HERIZI	Université de M'Sila	Examineur
Mr. Adbelhafid BENYOUNES	Université de M'Sila	Encadreur

Année universitaire : 2022 /2023

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Remerciement

Tout d'abord nous remercions ALLAH le tout puissant de nous donner la chance pour
réaliser notre travail de recherche

Nous tenons à remercier notre encadreur Mr. Benyounes Abdelhafid pour sa disponibilité,
ses orientations pertinentes, sans lequel, ce travail n'aurait pas vu le jour

Nous remercions également les membres de jury d'avoir acceptés de lire ce travail.

Ainsi nous tenons à exprimer nos sincères remerciements à notre chef de département et tous
les enseignants qui par leur compétences nous ont permis de poursuivre nos études

Nous présentons nos remerciements, nos respects et nos gratitude au groupe Sonalgaz qui
ont assuré le soutien affectif durant notre stage et l'intérêt qu'ils ont porté à la réussite de ce
modeste travail

Dédicace

Je dédie ce modeste travail :

À Mes parents :

Grâce à leurs tendres encouragements et leurs grands sacrifices, ils ont pu créer le climat affectueux et propice à la poursuite de mes études. Aucune dédicace ne pourrait exprimer mon respect, ma considération et mes profonds sentiments envers eux.

Je prie le bon Dieu de les bénir, de veiller sur eux, en espérant qu'ils seront toujours fiers de moi.

À ma petite et grande famille. À tous ceux qui m'ont aidé de près ou de loin. Ils vont trouver ici l'expression de mes sentiments de respect et de reconnaissance pour le soutien qu'ils n'ont cessé de me porter.

À tous mes amis et mes collègues de travail.

Saber MOUSSALI

Dédicace

Je dédie ce mémoire à deux personnes les plus chères dans ma vie : mes parents, pour leur amour spécial, leurs conseils et leur soutien durant mes années d'études qui nous consacré leur vie à bâtir la mienné, que ce travail soit

le fruit de vos innombrables sacrifices

À mes chers frères et sœurs qui sont toujours ma source de joie que je ne trouverais jamais assez de mots pour leurs exprimer mon amour

Et à tous ceux qui me sont chers de prêt ou de loin

Chaima SALEM

Résumé

Dans ce travail on s'intéresse à l'application des algorithmes d'apprentissage automatiques pour la détection et la classification des défauts dans les turbines à gaz. L'apprentissage automatique est une méthode d'analyse des données qui automatise la construction de modèles analytiques. Il s'agit d'une branche de l'intelligence artificielle basée sur l'idée que les algorithmes peuvent apprendre à partir des données, détecter, identifier et prendre des décisions. Grâce à leurs efficacités et rapidités, ces algorithmes nous donnent des résultats précis pour la détection et la classification des défauts donc un meilleur diagnostic des pannes dans un temps opportun.

Mot clef : Apprentissage automatiques, Turbine à gaz, Intelligence artificielle, détection des défauts et classification des défauts.

نولي اهتمام كبير في هذا العمل بتطبيق خوارزميات التعلم الآلي لكشف وتصنيف الأعطال في العنفات الغازية. التعلم الآلي هو طريقة لتحليل البيانات تعمل على اتمام بناء النماذج التحليلية. إنه فرع من فروع الذكاء الاصطناعي يعتمد على فكرة أن الخوارزميات يمكن أن تتعلم من البيانات وتكتشف وتتعرف وتتخذ القرارات. بفضل كفاءتها وسرعتها ، تعطينا هذه الخوارزميات نتائج دقيقة للكشف عن الأخطاء وتصنيفها ، وبالتالي تشخيص أفضل للأعطال في الوقت المناسب. **كلمات مفتاحية :** التعلم الآلي، العنفة الغازية، الذكاء الاصطناعي، الكشف عن الأخطاء، التصنيف.

Abstract

In this work we are interested in the application of machine learning algorithms for the detection and classification of faults in gas turbines. Machine learning is a data analysis method that automates the construction of analytical models. It is a branch of artificial intelligence based on the idea that algorithms can learn from data, detect, identify and make decisions. Thanks to their efficiency and speed, these algorithms give us precise results for the detection and classification of faults, therefore a better diagnosis of faults in a timely manner.

Keywords : Machine Learning, Gas Turbine, Artificial Intelligence, classification.

Table des Matières

Introduction Générale	1
------------------------------------	---

CHAPITER I : GENERALITES SUR LES TURBINE A GAZ

I.1. Introduction	3
I.2. Présentation de L'unité turbine à gaz de M'sila	3
I.3. Description de la central M'SILA III	4
I.3.1. Le Compresseur.....	5
I.3.2. Chambre de combustion.....	6
I.3.3. Les brûleurs de la chambre de combustion	8
I.3.4. La turbine.....	8
I.3.5. L'échappement.....	9
I.4. Principe de fonctionnement d'une Turbine à Gaz.....	10
I.5. Diagnostique	12
I.6. Définitions et terminologie.....	12
I.7. Les Défauts Des Turbines À Gaz.....	14
I.7.1. Défaut de vibration.....	14
I.7.1.1 Désalignement	15
I.7.1.2 Déséquilibre (Défaut de Balourd)	16
I.7.2. Défaut de la Haute Température à l'Echappement.....	17
I.7.3. Défaut de combustion.....	17
I.7.4. Défaut du Surcharge Compresseur	18
I.8. Détection des défauts.....	18
I.8.1. Détection des Défaut de la Haute Température à l'Echappement et le Défaut de Combustion.....	19
I.8.2. Détection du Défaut de surcharge compresseur.....	20
I.8.3. Détection des défauts de vibrations.....	21
I.8.3.1 Définition D'une Vibration.....	22
I.8.3.2 Caractéristiques Dunne Vibration.....	22
I.8.3.3 Surveillance des vibrations.....	23
I.9. Conclusion.....	24

CHAPITER II : INTRODUCTION A L'APPRENTISSAGE AUTOMATIQUE

II.1. Introduction	25
II.2 Intelligence Artificielle.....	25
II.2.1 Historique.....	26
II.2.2 L'apprentissage automatique (Machine learning).....	27
II.2.3 Pourquoi utiliser l'apprentissage automatique ?	28
II.3. Types d'apprentissage automatique.....	29

II.3.1. Apprentissage supervise.....	29
II.3.1.1 Les étapes de l'apprentissage supervisé.....	29
II.3.1.2 Les techniques d'apprentissage supervisé.....	30
1. Régression linéaire.....	30
2. SVM (machine à vecteur de support).....	31
3. Arbre de décision	33
4. K- plus proches voisins.....	33
5. Forêts aléatoires	34
6. Réseau neurones artificiel.....	34
7. Adaboost.....	34
8. XGBoost.....	35
9. Gradient Boosting.....	36
10. Naive Bayes	37
II.3.2. Apprentissage non supervisé.....	37
II.3.2.1 Types de problèmes d'apprentissage non supervisé.....	38
II.3.2.2 Applications de l'apprentissage non supervisé.....	38
1. Détection d'anomalies dans la finance.....	39
2. Clustering	39
3. Moteurs de recommandations et publicité personnalisées.....	39
II.3.2.3 Les techniques d'apprentissage non supervisé.....	39
1. Méthode des K-Means.....	39
2. Algorithme à priori.....	40
II.3.3. Apprentissage par renforcement	41
II.3.3.1 Les étapes de l'apprentissage par renforcement	42
II.3.3.2-Les techniques d'apprentissage par renforcement.....	42
1. Q-Learning.....	42
2. Deep Q- networks.....	42
II.3.3.3. Évaluation et sélection de modèle	43
a. Courbe ROC.....	43
b. Rappel.....	44
c. Précision.....	44
d. F-mesure	44
e. La validation croisée.....	45
f. La régularisation.....	46
g. Accuracy	46
II.4 Conclusion	46
CHAPTER III : RESULTATS ET DISCUSSIONS	
III.1. Introduction	47
III.2 Diagnostic des défauts des vibrations en vigueur	47
III.3. Critères D'évaluation.....	48
III.4. Environnement de Travail	50

III.5. Collection et Prétraitement du DATA.....	50
III.5.1. Kurtosis ;.....	51
III.5.2. Skewness.....	51
III.5.3. Facteur de forme.....	52
III.5.4. Impulsion	52
III.5.5. RMS	52
III.5.6. Moyenne absolue	52
III.5.7. Écart type	52
III.5.8. Facteur de dégagement.....	53
III.5.9. Entropie	53
III.5.10. Peak to Peak.....	53
III.5.11. Facteur de crête.....	53
III.6. Phase d'apprentissage.....	55
III.7 Résultats de simulation	55
III.7.1 Résultats d'apprentissage des Modelés	55
1. Le Réseau de Neurone Multicouches (MLP)	55
2. Modèle "Arbre de décision"	56
3. Modèle "Random Forest"	57
4. Modèle "Linear SVM"	57
5. Modèle "KNeighbors"	58
6. Modèle "Naive Bayes"	58
7. Modèle "XGboost"	59
8. Modèle "Adaboost"	59
III.7.2 Evaluation des Modèles	59
III.7.3 Etude comparative et Choix d'Algorithme	62
III.8. Conclusion.....	62
Conclusion générale.....	62

Liste de Figures

Figure I.1 : La Centrale de M'sila III.....	3
Figure I.2 : Turbocompresseur V94.3A de la centrale de M'sila.....	4
Figure I.3 : Compresseur (15) étages de la turbine V 94. 3A.....	5
Figure I.4 : Schéma présentatif d'une chambre de combustion.....	6
Figure I.5 : Chambre de combustion.....	6
Figure I.6 : brûleur hybride à deux combustibles.....	7
Figure I.7 : Les 4 étages mobiles de la turbine.....	8
Figure I.8 : Corps d'un Echappement d'une Turbine à Gaz.....	9
Figure I.9 : Principe de fonctionnement de turbina a gaz.....	10
Figure I.10 : Etape de diagnostic d'un Défaut.....	12
Figure I.11 : Défauts d'alignement d'arbres.....	13
Figure I.12 : Désalignement sur le plan vertical et horizontal.....	14
Figure I.13 : conséquence de désalignement sur le rotor d'une turbine.....	15
Figure I.14 : conséquence d'un balourd sur un rotor d'une turbine.....	16
Figure I.15 : Réparation d'un échappement d'une turbine à gaz.....	16
Figure I.16 : Dégradation de la chambre de combustion suite au défaut de combustion (TAG V94.3A4 de la centrale de M'sila).....	17
Figure I.17 : Défaillance des aubes compresseur (TAG V94.3A de M'sila).....	17
Figure I.18 : Points de mesure des vibrations (absolu et relative).....	21
Figure I.19. Caractéristiques d'une vibration Sinusoïdale ou quelconque.....	22
Figure I.20. Racks de surveillance des vibrations On-line.....	23
Figure I.21. Surveillance des vibrations Off-line.....	23
Figure I.22. Outils de collection des données Off-line.....	23
Figure II.1 : Sous-domaines of AI.....	25
Figure II.2 : AI Applications.....	26
Figure II.3 : L'approche traditionnelle.....	28
Figure II.4 : Approche d'apprentissage automatique.....	28
Figure II.5 : L'apprentissage automatique peut aider les humains à apprendre.....	28
Figure II.6 : Exemple de régression linéaire simple.....	30
Figure II.7: Data analytic post.....	31
Figure II.8: Classification of data by support.....	32
Figure II.9: Noyau.....	32
Figure II.10: Un arbre de décision.....	33
Figure II.11: K-plus proches voisins (apprentissage paresseux).....	34
Figure II.12 : algorithme Adaboost.....	35
Figure II.13: Algorithme XGboost.....	35
Figure II.14: Gradien boosted trees pour la régression.....	36
Figure II.15 : Algorithme Naive Bayes.....	37
Figure II.16: Apprentissage en profondeur clustering non supervisé.....	38
Figure II.17: Treillis d'itemset fréquents.....	39

Figure II.18 : Un algorithme de détection d'anomalies.....	39
Figure II.19 : La méthode des K-moyennes, ou K-means clustering.....	41
Figure II.20 : Reinforcement Learning Components.....	42
Figure II.21 : Un réseau de neurones mappant un état d'entrée à sa paire correspondante (action, valeur q) (Image de l'auteur)	44
Figure II.22 : La courbe ROC de trois estimateurs d'épitope.....	44
Figure III.1. Décomposition en FFT d'un signal de vibration	50
Figure III.2. Seuil de Jugement de vibration	51
Figure III.3. Chaîne de mesure et collecte des données	56
Figure III.4. Signal de vibration des données collectées.....	56
Figure III.5. Représentation graphique des indicateurs	55
Figure III.6. Résultat d'apprentissage pour le Modèle 'MPL'	56
Figure III.7. Résultat d'apprentissage pour le Modèle 'Arbre de décision'	56
Figure III.8 Résultat d'apprentissage pour le Modèle 'Random Forest'	57
Figure III.9 Résultat d'apprentissage pour le Modèle 'Linear SVM'	57
Figure III.10 Résultat d'apprentissage pour le Modèle 'KNeighbors'	58
Figure III.11 Résultat d'apprentissage pour le Modèle 'Naive Bayes'	58
Figure III.12 Résultat d'apprentissage pour le Modèle 'XgBoost'	59
Figure III.13 Résultat d'apprentissage pour le Modèle 'AdaBoost'	59
Figure III.14 Résultats d'évaluation des Modèle	61

Symboles

Indice	Mot clé	Unité
P	Pression	[b]
T	Temperature	[C°]
I.G.V	Inlet Gas Valve	/
H.T.E	Haute Temperature à l'échappement	/
B	Induction magnétique	[T],[mT]
V	Vitesse	[m/s]
Ac	Amplitude crête	/
Acc	Amplitude crête à crête	/
A _{eff}	Amplitude efficace	/
<i>f</i>	Fréquence	[Hz]
T	Période	[s]
IA	Intelligence Artificielle	/
ML	Machine Learning	/
SVM	machine à vecteur de support	/
FFT	Fast Fourier Transform	/
TAG	Turbine à Gaz	/

INTRODUCTION GENERALE

Les turbines à gaz sont largement utilisées dans diverses industries en raison de leur grande efficacité, de leur conception, compacte et de leur polyvalence. Elles jouent un rôle essentiel dans la production d'énergie, l'aviation et d'autres applications où une conversion d'énergie fiable et efficace est cruciale. Cependant, l'apparition de défauts dans les turbines à gaz peut compromettre leurs performances, augmenter les coûts de maintenance et présenter des risques pour la sécurité. Par conséquent, la détection précoce et la classification de ces défauts sont essentielles pour garantir le fonctionnement continu et optimal des turbines à gaz. Les différents aspects de surveillance prennent plus d'importance, il est essentiel de détecter les défauts le plus rapidement possible, afin d'éviter une aggravation du dysfonctionnement et donc une panne des processus; qui selon l'application peut résulter en des pertes monétaires, troubles environnementaux voire sanitaires, [1].

Les vibrations, mauvaise combustion et la haute température à l'échappement sont considérés les défauts les plus fréquents et plus connus au monde des turbines à gaz et son diagnostic tardif de ces défauts constitue un danger imminent aux conséquences particulièrement graves notamment le problème de vibration.

Ces dernières années, l'intelligence artificielle (IA) et les techniques d'apprentissage automatique ont émergé comme des outils puissants pour la détection et la classification des défauts dans des systèmes complexes. Ces technologies ont révolutionné la manière dont nous analysons les données, extrayons des informations et prenons des décisions éclairées. En exploitant les capacités de l'IA et de l'apprentissage automatique, il devient possible de développer des systèmes de détection de défauts robustes et précis, capables d'aider à l'identification précoce des défauts, ce qui permet une maintenance proactive et une réduction des temps d'arrêt.

Ce Mémoire vise à explorer l'application des techniques d'apprentissage automatique pour la détection et la classification des défauts de la turbine à gaz. Le mémoire se compose de trois chapitres qui approfondissent les différents aspects de cette recherche

Dans le premier chapitre, nous fournirons une description détaillée des turbines à gaz et de leur fonctionnement. Nous examinerons en détail les différents composants des turbines à gaz, leur rôle et leur interaction dans le processus de conversion de l'énergie. Nous explorerons également les avantages et l'importance des turbines à gaz dans la production d'énergie, l'aviation et d'autres applications industrielles. Comprendre le fonctionnement des turbines à gaz est essentiel pour

identifier les défauts potentiels et développer des systèmes de détection et de classification des défauts efficaces.

Le chapitre 2 présentera une introduction aux algorithmes d'apprentissage automatique et leur importance dans la détection et la classification des défauts. Nous explorerons différentes techniques telles que l'apprentissage supervisé, l'apprentissage non supervisé et l'apprentissage par renforcement. De plus, nous discuterons des avantages et des limitations des différents algorithmes, notamment les arbres de décision, les machines à vecteurs de support(SVM), les réseaux neuronaux et les méthodes d'ensemble. Nous mettrons en évidence la puissance et la polyvalence des algorithmes d'apprentissage automatique dans l'analyse de schémas de données complexes, la détection des anomalies et la prise de prédictions précises. Ce chapitre posera les bases pour l'application de ces algorithmes dans la détection des défauts des turbines à gaz.

Le chapitre 3 présentera la méthodologie utilisée pour détecter et classer les défauts dans les turbines à gaz en utilisant des algorithmes d'apprentissage automatique. Nous décrirons en détail la configuration expérimentale, y compris la collecte des données, les techniques de prétraitement et l'ingénierie des caractéristiques. De plus, nous expliquerons la mise en œuvre des modèles d'apprentissage automatique sélectionnés et les métriques d'évaluation utilisées pour évaluer leur performance. Les résultats obtenus en appliquant ces algorithmes aux données réelles des turbines à gaz seront présentés et analysés, démontrant ainsi l'efficacité et la puissance de l'IA et de l'apprentissage automatique dans la détection et la classification des défauts.

Et finalement nous résumerons les principales conclusions et contributions de ce mémoire. Nous mettrons l'accent sur l'importance de tirer parti du pouvoir de l'intelligence artificielle et de l'apprentissage automatique dans la détection et la classification des défauts des turbines à gaz. La capacité de ces technologies avancées à analyser de vastes quantités de données, à détecter des schémas subtils et à faire des prédictions précises a révolutionné le domaine du diagnostic des défauts. Nous discuterons des implications de cette recherche pour l'amélioration des pratiques de maintenance, la réduction des temps d'arrêt et l'amélioration des performances globales des turbines à gaz. De plus, nous fournirons des perspectives sur les défis et les opportunités à venir, suggérant des pistes de recherche futures dans ce domaine.

CHAPITRE I

GÉNÉRALITÉS SUR LES TURBINES À

I.1. Introduction

Jusqu'à aujourd'hui l'économie algérienne repose essentiellement sur l'énergie (gaz et pétrole) et pour produire de l'électricité, injecter les matières dans les zones de raffinage, de liquéfaction et de transport on utilise des machines appelées turbines. Les plus importantes et les plus utilisées sont les turbines à gaz. Une turbine à gaz, appelée aussi turbine à combustion, est une machine tournante thermodynamique appartenant à la famille des moteurs à combustion interne dont le rôle est de produire de l'énergie mécanique (rotation d'un arbre) à partir de l'énergie contenue dans un hydrocarbure (fuel, gaz...), [1]. Durant ces dernières années, les turbines à gaz industrielles ont joué un rôle important dans les systèmes de production de puissance, tels que les centrales nucléaires de puissance, les centrales de production d'électricité et les unités des hydrocarbures. Bien que ces équipements possèdent de nombreux avantages, vu leur haute sensibilité à l'influence de variation de la température de l'air ambiant, qui change considérablement entre le jour et la nuit, l'été et l'hiver, fait que le rendement thermique d'exploitation de ces machines se trouve affecté,[1].

La turbine à gaz comme toute installation ou équipement, se détériore dans le temps sous l'action de causes multiples surcharge en cours de fonctionnement, mauvaise exploitation, action des agents corrosifs, chimique, atmosphérique,... etc. Ces détériorations peuvent être à l'origine des arrêts de fonctionnement (pannes) de la turbine, des perturbations considérables de production, de l'augmentation des coûts de production de façon générale. Pour cela il est indispensable de maitre un plan de maintenance préventive afin de diminuer ces pannes et de minimiser ces couts.

I.2. Présentation de L'unité turbine à gaz de M'sila :

L'unité de M'sila est une ensemble des centrales thermiques basées sur les turbines à gaz, se compose de quatre centrales de production de l'électricité, la première s'appelle M'sila I est composée de deux groupes de turbine à gaz de type 5001P (CONSTRUCTEUR THOMASSEN) de puissance unitaire 25 MW aux bornes de l'alternateur. La deuxième M'sila II est composée de trois groupes de turbines à gaz de type 9001E (CONSTRUCTEUR ALSTHOM) chacun de puissance nominale 110 MW, la troisième M'sila III est composée de deux groupes de turbines à gaz de type V94.3A4 (CONSTRUCTEUR ANSALDO ENERGIA) chacun de puissance nominale 215 MW, et la dernière BBA est composée de 02 groupes de 20 MW PWPS (pratt and whitney), alors la puissance totale installée est de 850 MW, aux bornes de l'usine BU. L'évacuation de l'énergie produite de chaque centrale, vers le poste HT, est assurée par un transformateur principal dimensionné pour transférer la puissance de l'alternateur.

La liaison entre les bornes du générateur et les bornes coté basse tension du transformateur principal est réalisé par gaine coaxiales. En ligne sur les gaines coaxiales est prévu le disjoncteur du générateur.

I.3. Description de la central M'SILA III

La centrale turbine à gaz M'SILA III est conçue et réalisée en 2009 dans le but de sécuriser la région EST en matière d'alimentation en énergie électrique, elle a été mise en service en 2009 par le constructeur italien ANSALDO ENERGIA pour renforcer une ancienne installation de deux centrales turbine à gaz M'SILA I, 2*25 MW et M'SILA II 3*110MW toujours sont opérationnelles en raison même de qualités de son services d'entretien et de maintenance, cette centrale occupe avec son bâtiment administratif, (voir la **Figure I.1**) ses ateliers, ses installations anti-incendie, les aéro-réfrigérants, des alternateurs et les bacs à fuel, un site de trois hectares. Ses équipements de dernière génération deux turbines ANSALDO ENERGIA de 215 MW et un poste d'évaluation de 220 KV chacun lui assurent une polyvalence et un fonctionnement en mode de combustion hybride : le gaz comme combustible principal et le fuel comme combustibles de secours.



Figure I.1. La Centrale de M'sila III.

La centrale M'SILA III, contient deux groupes TG pour la production d'énergie électrique identiques qui présentent chacune une puissance de 215MW condition ISO et une tension de 225 KV par l'intermédiaire d'un transformateur élévateur de tension vers le réseau de transport.

La turbine à gaz est composée de trois éléments principaux qui sont :

- Compresseur axial.
- Chambre de combustion.
- Turbine.

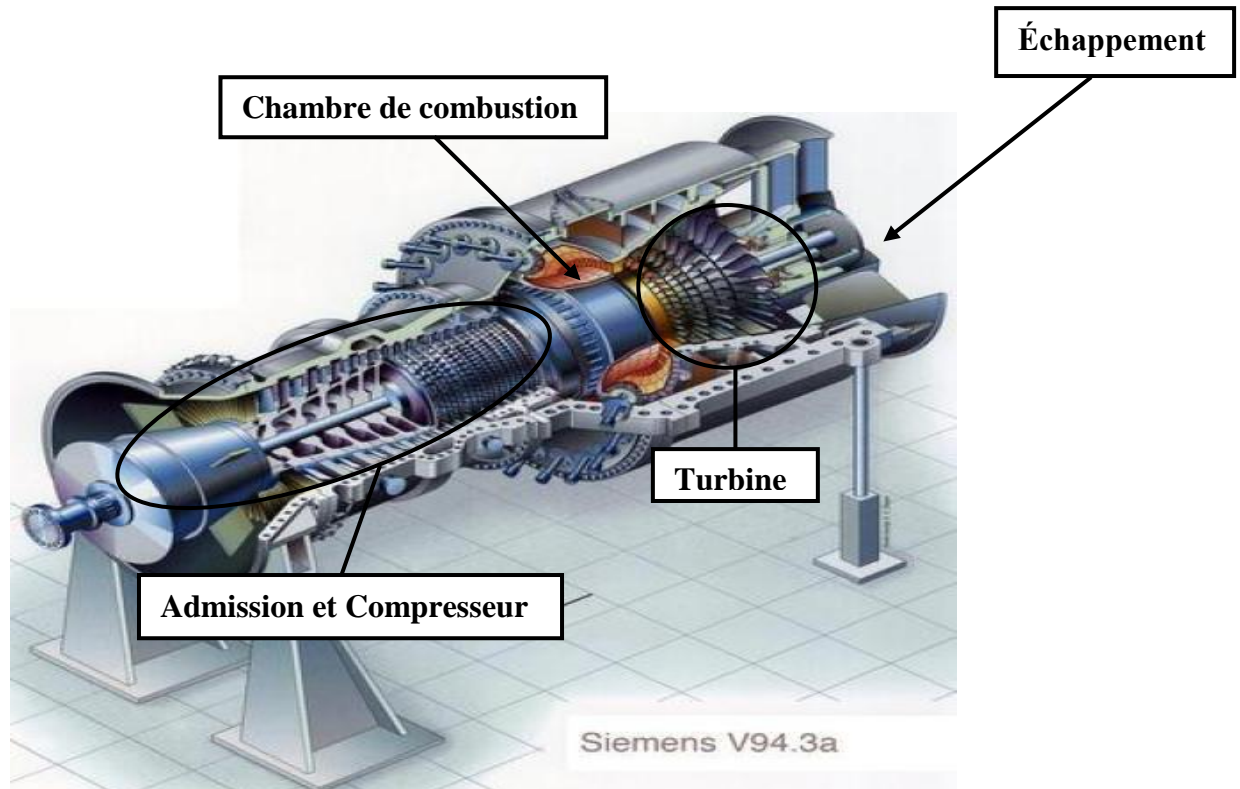


Figure I.2. Turbocompresseur V94.3A de la centrale de M'sila[2].

I.3.1. Le Compresseur :

Le compresseur a comme mission principale l'alimentation en air sous pression de la chambre de combustion de la turbine à gaz pour son mélange avec le combustible. Le compresseur utilisé dans la turbine V94.3A est de type axial à 15 étages. Il aspire l'air à la pression atmosphérique et le refoule à une pression d'environ 17 bars selon le régime de fonctionnement et les besoins requis, ce qui donne un taux de compression de 17 bars, [2].

Chaque étage du compresseur est formé par :

- Un aubage rotorique, disposé en couronne sur l'arbre.
- Un aubage statorique, sur la face interne de l'enveloppe extérieure.

Le compresseur dispose en amont de son entrée un dispositif de filtration de l'air, constitué d'un préfiltre et de trois filtres, qui permettent tous de retenir la plupart des résidus et particules en suspension dans l'air. La propreté de l'air admis dans le compresseur est d'une importance primordiale

pour sa longévité et la bonne marche du groupe. De l'air destiné au refroidissement de la turbine est également soutiré au niveau de plusieurs étages du compresseur.

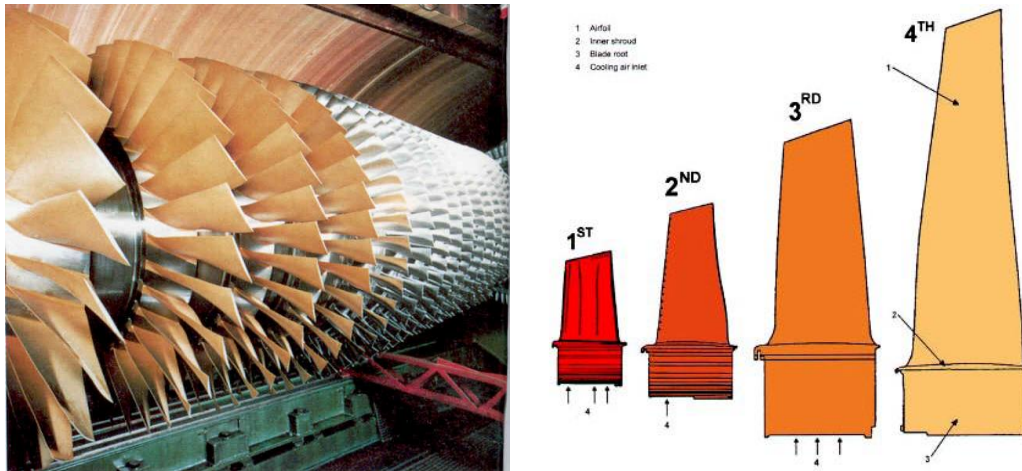


Figure I.3 : Compresseur (15) étages de la turbine V 94. 3A, [2].

I.3.2. Chambre de combustion

La chambre de combustion est l'une des trois parties d'une turbine à gaz dans sa forme simple. C'est la zone dans laquelle un combustible gazeux ou liquide est injecté sous pression, puis brûlé avec l'air comprimé, avec un fort excès d'air afin de limiter la température des gaz d'échappement. Le rôle de la chambre de combustion est de convertir l'énergie chimique contenue dans le combustible en énergie thermique. Cette énergie peut ainsi être augmentée à travers la turbine pour produire la puissance requise pour entrainer le compresseur ainsi que celle nécessaire pour entrainer une charge désirée (générateur par exemple).

Pour jouer ce rôle, la chambre de combustion sert à :

- Mélanger l'air et le combustible ;
- Allumer le mélange air-combustible grâce à des bougies;
- Contenir le mélange durant la combustion,
- Assuré la répartition de la température des gaz chauds à la sortie.

La chambre de combustion comprend plusieurs compartiments à savoir : une zone primaire, une zone secondaire, une zone de dilution, un ensemble varié de gicleurs et dispositif de gestion de transfert de chaleur.

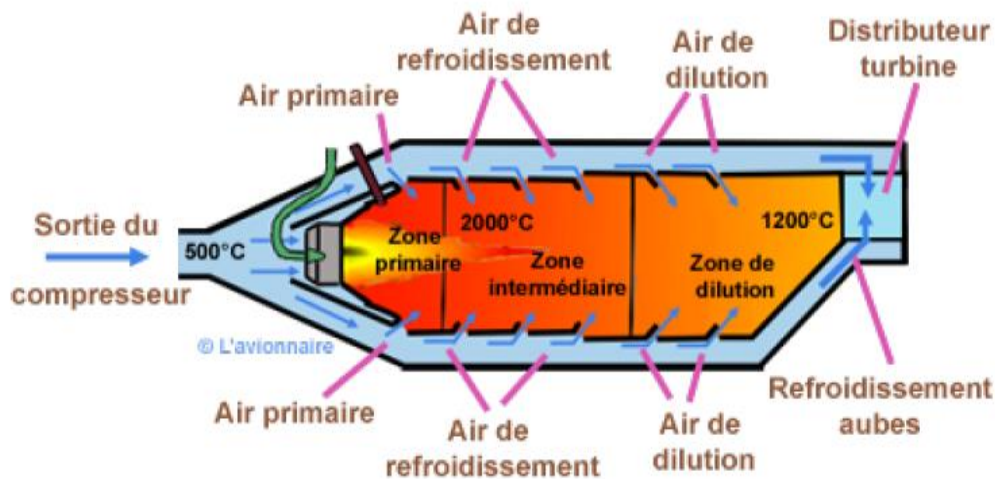


Fig. I.4: Schéma présentatif d'une chambre de combustion,[2].

La turbine à gaz ANSALDO V94.3A est dotée d'une chambre de combustion de type annulaire équipée de vingt-quatre brûleurs hybrides répartis régulièrement sur son périmètre, afin d'assurer une zone de température homogène; tous les brûleurs sont constitués de plusieurs injecteurs. Ces brûleurs permettent de réduire les émissions de NO par un procédé sec tout en maintenant les émissions de CO à un niveau minimal. Le carburant est amené à la chambre de combustion sous très forte pression par des pompes, il est alors vaporisé par les injecteurs. Les gouttelettes ainsi formées doivent être les plus fines possible afin de pouvoir réduire la longueur de la chambre de combustion.

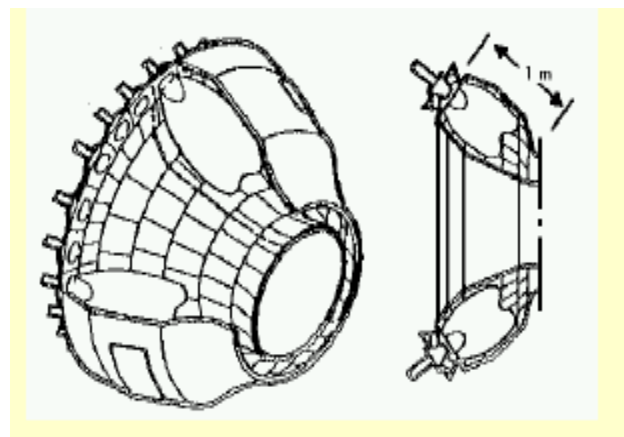


Figure I.5: Chambre de combustion

I.3.3. Les brûleurs de la chambre de combustion

Les brûleurs hybrides fonctionnent à double combustible (gaz naturel et liquide, fuel). Ces brûleurs ont pour but de :

- permettre le réglage des débits air, gaz, pour obtenir de meilleures conditions de combustion telle que les conditions stœchiométriques.
- donner à la flamme des dimensions appropriées à celle de l'enceinte où doit s'effectuer la combustion.
- réaliser le mélange de l'air et du gaz.

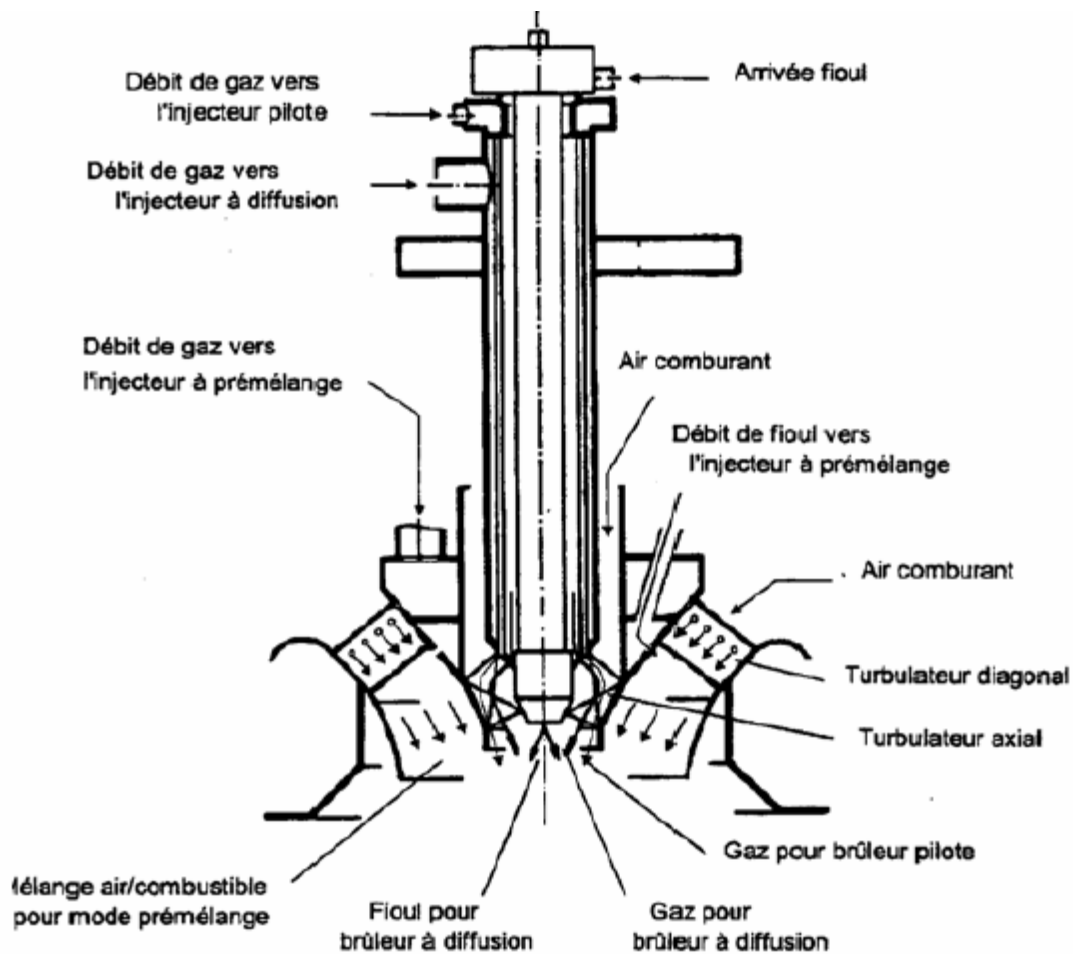


Figure I.6 : brûleur hybride à deux combustibles, [2].

I.3.4. La turbine

La turbine de détente est constituée de 4 étages au travers desquels se détendent les gaz chauds issus de la chambre de combustion. Comme pour le compresseur, chaque étage de la turbine est constitué d'une paire de couronnes d'aubes ; l'une statorique et l'autre rotorique, à la différence qu'ici, ce ne sont pas les aubes rotoriques qui refoulent les gaz circulant, mais plutôt le passage de ceux-ci, en se détendant au travers de chaque étage qui amorcent le mouvement du rotor, permettant la conversion

de l'énergie de pression des gaz brûlés, en énergie cinétique sur le rotor. Les aubes ont un recouvrement qui les protège contre la corrosion, l'oxydation et la déformation. Les gaz entrent aux travers des aubes avec une température très importante 1300°C et sortent de la turbine avec une température de 590°C et à une pression d'environ 1 bar, le débit des gaz d'échappement est d'environ 655 kg/s . En raison de la forte exposition de la turbine de détente et de ses organes aux hautes températures, son refroidissement est d'une importance capitale pour garantir la fiabilité et le fonctionnement optimale de la turbine, [2]. Son principe de fonctionnement se base sur des soutirages d'air aux étages appropriés du compresseur, évitant ainsi de faire appel à un système de refroidissement externe. Les différents étages d'aubes de la turbine nécessitent chacun de l'air de refroidissement à une pression déterminée, c'est pourquoi il est procédé à des soutirages à différents étages du compresseur. Ainsi chaque étage de la turbine est refroidit comme suit :

- Etage 1 : air soutiré de l'échappement du compresseur.
- Etage 2 : air soutiré du 15^{ème} étage du compresseur pour les aubes fixes, et en aval du 14^{ème} étage pour les aubes mobiles.
- Etage 3 : air soutiré du 10^{ème} étage du compresseur pour les aubes fixes, et en aval du 14^{ème} étage pour les aubes mobiles.
- Etage 4 : air soutiré du 5^{ème} étage pour les aubes fixes, et du 14^{ème} étage pour les aubes mobiles.



Figure I.7: Les 4 étages mobiles de la turbine

I.3.5.L'échappement

Les gaz issus de turbines à gaz quittent le dernier étage de détente avec une très grande vitesse. Il est nécessaire de réduire considérablement cette vitesse avant de rejeter les gaz à l'atmosphère de façon à traiter convenablement le bruit à l'échappement de la turbine.

Les gaz d'échappements traversent le cadre d'échappement et le diffuseur, constitué d'une série de déflecteur transformant la direction axiale des gaz en direction radiale tout en minimisant les pertes par frottement les gaz parvient ensuite au caisson d'échappement et sont évacués à l'atmosphère par le système d'échappement.

Les silencieux permettant de réduire le bruit de l'échappement des gaz sont généralement insérés sur des conduits d'échappement de section polygonale, par exemple rectangulaire, et notamment carrée.

De cette manière, les baffles évitant la propagation sonore sont plus simples à concevoir et présentent des parois plates. De plus, le rendement des silencieux est d'autant meilleur que la vitesse d'écoulement des gaz entre les baffles est homogène dans la section droite du conduit d'échappement.

Le diffuseur est inséré entre de la turbine et le silencieux. Il permet de réduire la vitesse des gaz en élargissant son écoulement dans la section du conduit à travers lequel s'échappent les gaz de la turbine. Le diffuseur assure donc la transition du flux d'un conduit à section circulaire vers un conduit à section polygonale.



Figure I.8: Corps d'un Echappement d'une Turbine à Gaz

I.4. Principe de fonctionnement d'une Turbine à Gaz

Dès que la ligne d'arbre est mise en mouvement par le moteur de lancement, l'air atmosphérique est aspiré, filtré et dirigé vers l'entrée du compresseur axiale au 15^{ème} étage. Pour prévenir le pompage du compresseur des vannes d'extraction d'air en aval de 13^{ème} étage « Vanne anti-pompage » sont en position ouverte pendant le démarrage, et les aubes à orientation variable (I.G.V) situées à l'entrée du compresseur sont en position dite « fermée ».

Lorsque la vitesse de rotation atteinte 95% de sa valeur nominale, un relai de vitesse provoque la fermeture automatique des vannes d'extraction d'air du 13^{ème} étage et l'ouverture à une position prédéterminée des (I.G.V) situées à l'entrée du compresseur, à la sortie du compresseur axial, l'air

pénètre dans un espace annulaire entourant la chambre de combustion puis, dans l'espace situé entre l'enveloppe de chambre. Les brûleurs introduisent le combustible dans la chambre de combustion où il est mélangé à l'air de combustion venant du compresseur.

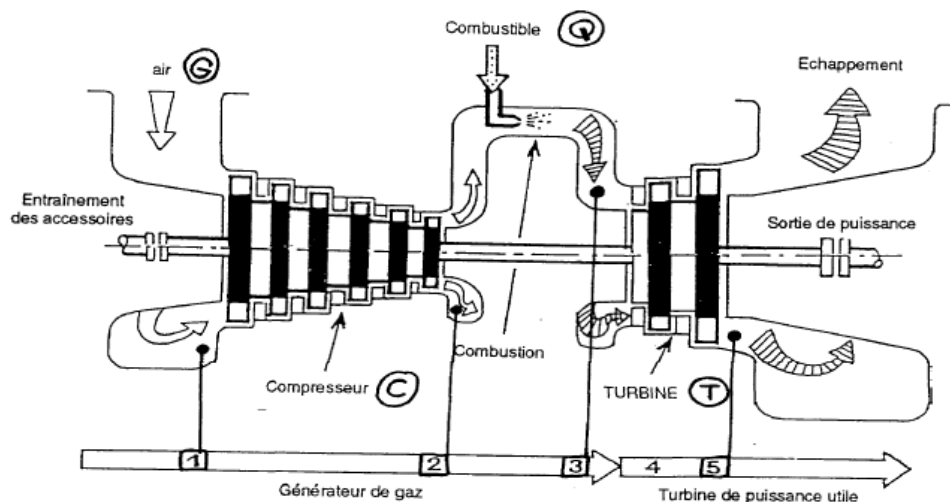
La mise à feu est réalisée par 24 bougies d'allumages (Mais une seule suffit pour réaliser la mise à feu).

Lorsque la turbine à presque atteint sa valeur nominale (3000 RPM par exemple) la pression des gaz à l'intérieur de chambre est suffisante pour provoquer le retrait des électrodes rétractables équipant les bougies d'allumage. Ainsi les électrodes sont protégées de l'action de la flamme.

Les gaz chauds provenant de la chambre de combustion attaquent les étages turbine, chaque étage est constitué par un ensemble des aubes fixes suivi d'une rangée d'aubes mobiles dans chaque rangée d'aubes fixe l'énergie cinétique de jet de gaz augmente parallèlement à la diminution de la pression dans la rangée adjacente d'aubes mobiles une partie d'énergie cinétique de jet et converti en travail utile transmis au rotor de la turbine.

Après leur passage dans les étages turbine, les gaz d'échappements traversent le cadre d'échappement et le diffuseur, constitué d'une série de déflecteur transformant la direction axiale des gaz en direction radiale tout en minimisant les pertes par frottement les gaz parvient ensuite au caisson d'échappement et sont évacués à l'atmosphère par le système d'échappement.

Le travail fourni au rotor de la turbine, une partie utilisée pour l'entraînement du compresseur axial et d'auxiliaires turbine, et l'autre sert à faire entrainer une charge désirée, [3].



I.9 : Principe de fonctionnement de turbina a gaz, [3].

I.5. Diagnostique

L'apparition d'un défaut conduit le plus souvent à un arrêt irrémédiable de la machine, en conséquence, un cout de réparation considérable pour l'entreprise et aussi une perte de production. Un bon diagnostic en temps opportun assuré la sécurité des personnes et des matériels. La démarche de

diagnostic consiste à trouver le défaut le plus plausible à partir des comportements observés et des informations suivantes :

Manifestations extérieures des défauts ou « symptômes » (fréquences ou contenu spectral, types d'évolution, localisation...), Contexte dans lequel apparaît le défaut, (conditions de fonctionnement, manœuvres d'exploitation, corrélations) et Connaissance de la machine, technologies favorisantes, historique d'incidents.

I.6. Définitions et terminologie

Ci-dessous, nous passons en revue certaines définitions spécifiques au domaine du diagnostic de panne utilisées dans ce mémoire :

Défaut

Un défaut est une déviation non permise d'au moins une propriété ou un paramètre caractéristique du système par rapport à un comportement nominal ou acceptable. C'est-à-dire tout écart entre la caractéristique observée et la caractéristique de référence. A titre d'information, une défaillance conduit à un défaut, mais un défaut n'induit pas nécessairement une défaillance.

Défaillance

Une défaillance est une altération ou cessation de l'aptitude d'un système à accomplir correctement sa (ses) fonction(s) requise(s) en assurant les performances définies dans des spécifications techniques. Sachant que cette défaillance peut être complète ou partielle.

Panne

Une panne est une interruption permanente de tout ou partie des capacités du système pour remplir sa fonction requise dans des conditions de fonctionnement spécifié.

Perturbation

Une perturbation est une entrée inconnue et incontrôlée qui affecte le système. Contrairement au défaut, qui est interne au système, la perturbation est une entrée externe.

Résidu

Un résidu est un signal conçu pour indiquer les anomalies fonctionnelles ou comportementales basé sur la différence entre les mesures et les calculs.

Surveillance

La surveillance d'un système est une tâche continue en temps réel permettant de déterminer l'état d'un système à travers l'enregistrement des informations pouvant indiquer la survenue d'éventuelles anomalies dans son comportement.

Supervision

La supervision consiste à prendre les décisions appropriées, pendant la phase de surveillance du système, afin de maintenir son fonctionnement nominal malgré l'apparition de défauts.

Diagnostic

Le diagnostic permet de déterminer le type, l'amplitude, la localisation et l'instant de détection des défauts. Il inclut les phases d'isolation et d'identification des défauts. Cependant, une autre définition donnée par Gertler, [4][5] en précisant que le diagnostic de défaut inclut également la première étape dite de détection des défauts. C'est aussi la définition adoptée dans ce mémoire.

En outre, beaucoup de revues et journaux scientifiques s'appuient sur l'abréviation FDI (Fault Detection and Isolation) pour décrire les trois phases détection, localisation et identification des défauts.

Détection de défaut

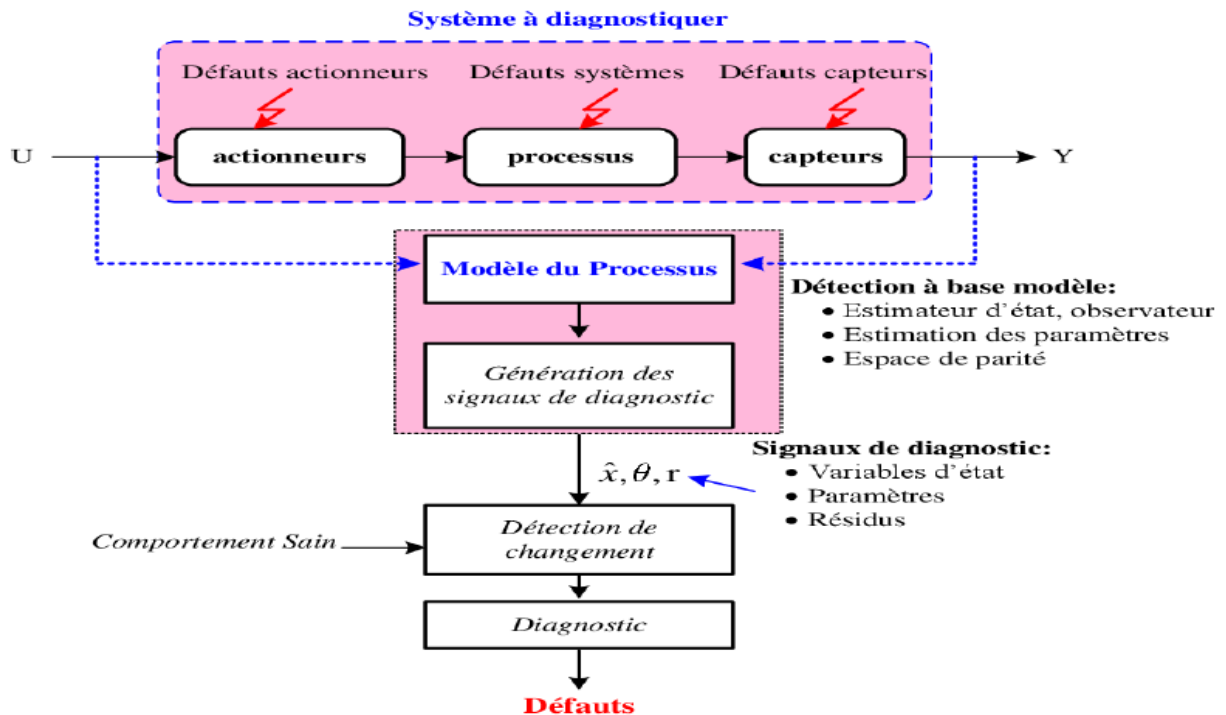
La détection de défaut détermine la présence ou non d'un défaut ainsi que l'instant de son occurrence.

Isolation de défaut

L'isolation de défaut est la détermination du défaut le plus vraisemblable permettant d'expliquer le changement du comportement du système, de sa localisation et de son instant d'occurrence.

Identification de défaut

L'identification de défaut est l'estimation de l'amplitude et l'évolution temporelle du défaut afin d'expliquer mieux le comportement du système. Elle est également considérée comme la dernière phase du diagnostic des défauts.



I.10 : Etape de diagnostic d'un Défaut, [6].

I.7. Les Défaits Des Turbines À Gaz

Les turbines à gaz sont soumises à plusieurs problèmes de leurs états dynamiques de fonctionnement entraînant des instabilités et des défauts, pouvant être dangereuses pour leur tenue mécanique provoquant des dysfonctionnements dans la totalité de l'installation de turbine. Il est donc essentiel de savoir tous les défauts qui peuvent influencer sur le bon état de la turbine pour maintenir le fonctionnement de ce type de machines tournantes en dehors de la zone d'instabilité.

I.7.1. Défaut de vibration

Toutes les machines en fonctionnement produisent des vibrations, images des efforts dynamiques engendrés par les pièces en mouvement. Ainsi, une machine neuve en excellent état de fonctionnement produit très peu de vibrations. La détérioration du fonctionnement conduit le plus souvent à un accroissement du niveau des vibrations. En observant l'évolution de ce niveau, il est par conséquent possible d'obtenir des informations très utiles sur l'état de la machine. Ces vibrations occupent une place privilégiée parmi les paramètres à prendre en considération pour effectuer un diagnostic. La modification de la vibration d'une machine constitue souvent la première manifestation physique d'une anomalie, cause potentielle de dégradations, voire de pannes, [6].

Le défaut de vibration se manifeste sur les paliers en plusieurs formes tel que :

I.7.1.1 Désalignement

Le défaut d'alignement est un phénomène mécanique, il apparait lors du mouvement de rotation de l'arbre avec les paliers. Deux types de désalignements existent : angulaire et parallèle, comme il est montré sur la **Figure I.11**.

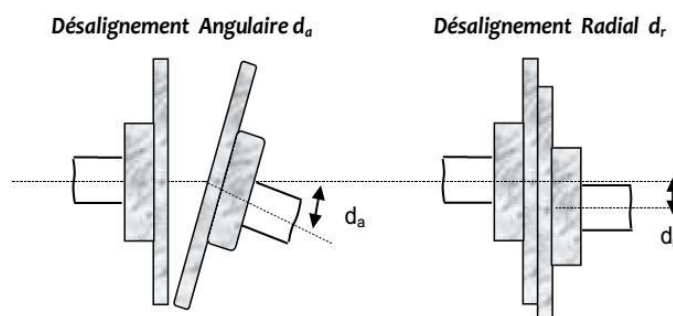


Figure I.11: Défauts d'alignement d'arbres, [6].

Le désalignement angulaire survient lorsque les axes des arbres se croisent. La vibration est radiale et axiale, un arbre courbé se comporte de façon identique. Lorsque le désalignement angulaire est critique, l'amplitude de la vibration axiale est supérieure à l'amplitude de la vibration radiale. Il

provoque des vibrations à la fréquence de rotation, ainsi qu'aux fréquences double, triple ou quadruple de la fréquence de rotation. La composante axiale de la vibration est particulièrement importante pour l'harmonie.

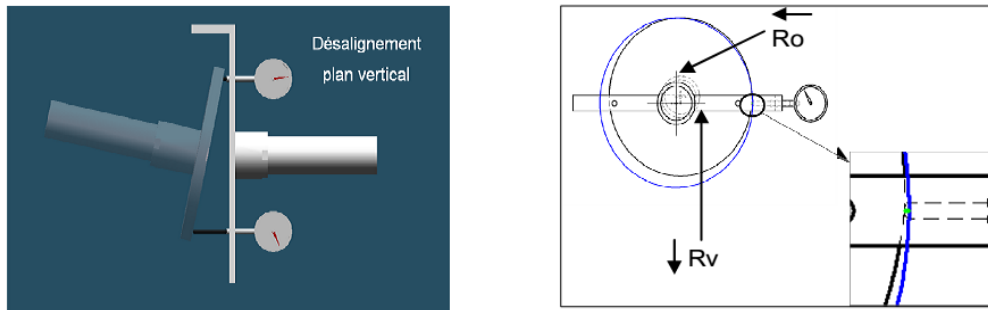


Figure I.12: Désalignement sur le plan vertical et horizontal, [6].

Le désalignement sur le plan vertical et horizontal peut affecter à la fois deux paliers du même arbre d'une machine ; les axes des deux paliers de même corps ne sont pas concentriques. Cette anomalie peut être la conséquence d'un défaut de montage d'un palier, mais également d'un mauvais calage des pattes de fixation ou d'une déformation de châssis (par exemple à la suite de contraintes thermiques), qui se traduit par une flexion de l'arbre du rotor, [7] [8].

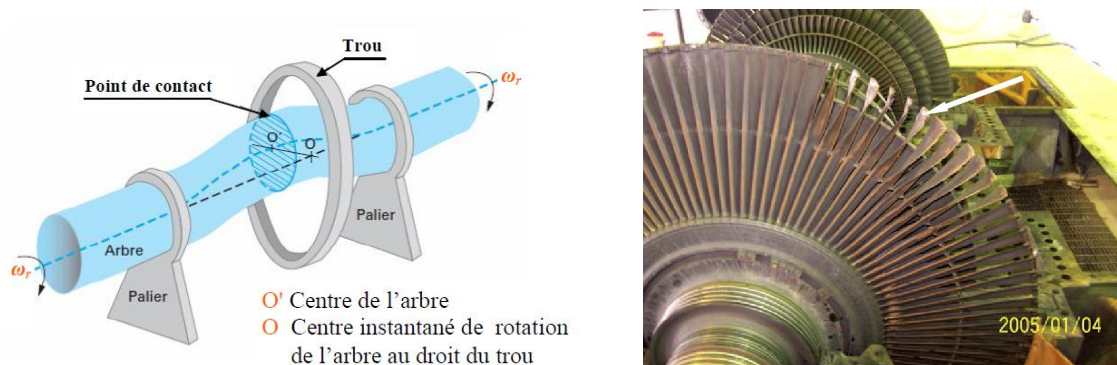


Figure I.13: conséquence de désalignement sur le rotor d'une turbine, [6].

I.7.1.2 Déséquilibre (Défaut de Balourd)

Le balourd est le résultat d'un ensemble de forces engendrées par les écarts existant entre les centres de gravité des différents éléments constituant le rotor et son axe de rotation. Cette force centrifuge, issue de ces défauts de « non-concentricité » ou d'ajustage, est proportionnelle à la masse de matière qui crée ce déséquilibre et au carré de la vitesse. Elle génère des efforts tournants qui vont agir sur les différents éléments constitutifs du rotor, les ancrages de la machine et les paliers jusqu'à

entraîner leur dégradation, [9]. Le défaut de balourd est l'une des principales causes de réduction de la durée de vie des équipements et il constitue l'un des principaux risques pour la sécurité. C'est pour cette raison sans doute qu'il fait l'objet de normes importantes, tant pour son évaluation que pour sa correction. Ainsi, pour une classe d'équilibrage donnée, comparer l'amplitude de la vibration induite à des seuils fixés dans des normes ou à des spécifications de constructeurs, permettra de statuer sur le caractère acceptable ou non du déséquilibre et sur la nécessité de procéder ou non à un équilibrage suivant une procédure bien définie, [10]. Le balourd, déséquilibre massique, peut provenir de défauts d'usinage, d'assemblage et de montage, mais aussi :

- a- **Balourd mécanique :** d'une altération mécanique (perte d'ailette, érosion ou encrassement...),
- b- **Balourd thermique :** d'une altération thermique ou d'une déformation suite à des dilatations différentes des matériaux constituant le rotor ou encore à des différences de température localisées.



Figure I.14: conséquence d'un balourd sur un rotor d'une turbine

I.7.2. Défaut de la Haute Température à l'Échappement

Les gaz d'échappement sortent à des températures élevées (de l'ordre de 500°C) d'où la relation générale entre température d'échappement et charge doit être observée et comparée aux données de références. Une haute température d'échappement sera une indication de détérioration des pièces intérieures, de fuite excessive ou d'un compresseur souillé. Elle peut aussi indiquer une surcharge (d'où l'intérêt de mesure de charge) ou dérèglement des dispositifs de commande.

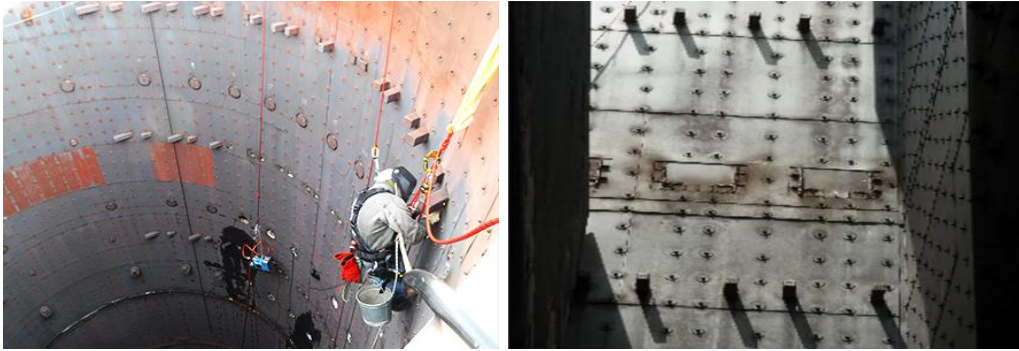


Figure I.15: Réparation d'un échappement d'une turbine à gaz

I.7.3. Défaut de combustion

L'instabilité de la combustion est généralement due à l'interaction du processus de combustion et de la pression de la flamme à l'intérieur de la chambre de combustion. Une certaine instabilité de la combustion est causée par des fluctuations de l'alimentation en air de la chambre de combustion, d'autres sont causées par des perturbations aérodynamiques générées par la chambre de combustion elle-même, et d'autres encore sont causées par des changements dans l'alimentation en combustible de la chambre de combustion, [11].



**Figure I.16: Dégradation de la chambre de combustion suite au défaut de combustion
(TAG V94.3A4 de la centrale de M'sila)**

I.7.4. Défaut du Surcharge Compresseur

Le phénomène de surcharge est un mode de fonctionnement instable, rencontré dans tous les systèmes de compression. La mesure de pression différentielle en amont du compresseur axial nous donne une bonne image du fonctionnement du compresseur et du taux de compression ainsi que un contrôle clair de l'état des aubes fixe et mobile.



Figure I.17: Défaillance des aubes compresseur (TAG V94.3A de M'sila)

I.8. Détection des défauts

La détection permet de découvrir tout écart du comportement anormal du système et alerte les opérateurs humains de supervision de la présence d'un défaut et son temps d'apparition. La détection de défauts fait l'objet de recherches depuis des décennies. Initialement, la recherche était axée sur l'aviation, mais a ensuite été appliquée à d'autres équipements tels que la turbine à gaz. La littérature sur ce sujet s'est enrichie grâce aux contributions de différents groupes de recherche, qui ont proposé différentes méthodes de détection de défauts en fonction de leur domaine d'expertise. Les améliorations apportées à ces méthodes ont également été améliorées en réponse aux exigences de l'industrie en matière de sécurité, de fiabilité et d'automatisation des systèmes.

La détection des défauts à un stade précoce offre la possibilité de planifier et de préparer les interventions nécessaires en les intégrant, si possible, dans les arrêts procédés ou en provoquant des arrêts programmés à des moments opportuns, en fonction notamment des impératifs de production.

I.8.1. Détection des Défaut de la H.T.E et le Défaut de Combustion

Le but du contrôle de la température d'échappement est de faire fonctionner la turbine à gaz avec une température d'entrée optimale dans presque toutes les conditions de charge de base. En effet, en charge de base, il est important de faire fonctionner la turbine à gaz à la température d'entrée adéquate : si la température d'entrée est trop élevée ceci pourrait provoquer des dommages et une contrainte thermique sur les composants chauds, si elle est trop faible, ceci signifie que la turbine à gaz n'est pas performante et qu'elle est sous-employée. Puisqu'il n'est pas possible de mesurer de façon précise la température d'entrée (en raison de valeurs très élevées), la température des gaz d'échappement est utilisée. La température d'entrée est liée à la température d'échappement par une corrélation impliquant la température d'entrée du compresseur et la vitesse de rotation. Par cette

corrélation, il est possible de faire fonctionner la turbine à gaz avec une température de gaz d'échappement constante et en même temps avec une température d'entrée de la turbine presque constante.

La température des gaz d'échappement (TET) est mesurée directement en aval de la turbine au moyen de 24 thermocouples triples qui sont situés le long de la circonférence du diffuseur de gaz d'échappement. Pour avoir une meilleure représentation de la température moyenne, la moyenne des 24 thermocouples peut être corrigée par une deuxième moyenne calculée au moyen de 6 thermocouples situés à la fin du conduit du diffuseur de gaz d'échappement. La valeur moyenne de la température d'échappement de la turbine est obtenue par la valeur moyenne des 24 thermocouples en cas de variations rapides de charge et par la valeur moyenne des 6 thermocouples en cas de fonctionnement de charge de base et de variations de charge lentes. Pour le contrôle de la TG, le canal A de chaque thermocouple est utilisé.

La distribution des températures des gaz d'échappement est utilisée afin d'évaluer les problèmes spécifiques qui peuvent se produire sur plusieurs brûleurs.

A la sortie de la turbine à gaz, les 24 thermocouples triples sont installés pour mesurer les changements rapides de température dans la zone chaude. Dans un but de contrôle (calcul de la température) la moyenne des 24 canaux 'a' est utilisée ; les canaux b et c sont utilisés dans des buts de protection. De plus, 6 thermocouples à un seul élément sont installés dans le conduit d'échappement, les signaux de ces thermocouples fournissent une valeur plus fiable de la température moyenne puisque les gaz chauds sont mélangés.

La distribution de la température des gaz d'échappement est contrôlée par le système de contrôle de la turbine à gaz, afin de détecter des anomalies dans la chambre de combustion. En effet, une distribution de la température non uniforme peut conduire à des contraintes thermiques sur le parcours du gaz chaud ou à une augmentation des émissions de NOx (à cause d'une augmentation de température locale).

L'évaluation de la distribution de température est basée sur la déviance d'un thermocouple simple par rapport à la température moyenne, en cas de déviation importante, une alarme se déclenche.

La distribution de température est affichée sur le panneau de contrôle de l'opérateur ; il est recommandé de vérifier régulièrement la distribution. Pendant le fonctionnement normal (à savoir à l'exception du démarrage ou de la commutation) l'on s'attend à avoir une déviation maximum de 50°K entre la valeur maximale et la valeur minimale des thermocouples (pas à partir de la moyenne). En cas de déviations supérieures, il est recommandé de rechercher la cause de cette augmentation.

La détection des défauts de la haute température à l'échappement se fait par la comparaison entre la température d'échappement mesurée et une température de consigne déclarée sur le système de contrôle commande.

I.8.2. Détection du Défaut de surcharge compresseur

La pression en amont du compresseur est indiquée par un transducteur de pression. En plus de l'élément de mesure électrique, ce dispositif possède un afficheur optique pour lecture directe. Le signal provenant de ce transducteur de pression est aussi utilisé pour mesurer le taux de compression du compresseur. Le taux de compression à la sortie du compresseur par rapport à la pression d'entrée du compresseur est envoyé au contrôleur de la limite du taux de compression du compresseur. Ce contrôleur évite un taux de compression excessivement élevé du compresseur, ce qui provoquerait le pompage du compresseur.

La chute de pression entre le conduit d'entrée et l'entrée du compresseur (immédiatement en amont des aubes directrices à pas variable d'entrée) est mesurée par trois pressostats. Les pressostats différentiels sont raccordés en parallèle et débutent la procédure de panne de la turbine à gaz en appliquant la logique deux sur trois, si la pression différentielle descend au-dessous d'une valeur donnée lorsque la vitesse de la turbine est au-dessus 84 % de la vitesse nominale, ce qui arrive lorsque le compresseur oscille.

I.8.3 Détection des défauts de vibrations

La détection des défauts des vibrations se fait à travers l'analyse des signaux délivrés par des capteurs de vibrations installés sur les paliers, on distingue trois types de capteurs de vibrations utilisés actuellement en industrie pour la maintenance prédictive sont les capteurs de déplacement, les capteurs inductifs de vitesse, et les accéléromètres ces capteurs transforment les vibrations aux points de mesure en une tension électrique.

a- Le Capteur de Déplacement : Le capteur de proximité (de déplacement) permet de mesurer les vibrations relatives de l'arbre d'une machine tournante. Il est également utilisé pour mesurer d'autres paramètres tels que la position axiale et la référence de phase. Il fournit une tension proportionnelle à la distance entre la tête du capteur et la cible.

c- Le capteur de vitesse – Le Vélocimètre : C'est un capteur sismique de type électromagnétique. Il se compose d'un aimant permanent solidaire du boîtier qui génère une induction magnétique constante **B**. et d'une bobine mobile liée au boîtier par deux ressorts plans qui peut se déplacer dans la direction des lignes de champ. Sous l'effet de la vibration, une force

électromotrice (f.e.m) d'induction e est alors générée, qui est le produit de la vitesse v de vibration, de l'induction magnétique B et la longueur l de l'enroulement.

$$e = B.l.v \dots\dots\dots I.1$$

et pour B et l constante ($B.l = k$)

$$e = k.v \dots\dots\dots I.2$$

La f.e.m d'induction est proportionnelle à la vitesse de vibration.

c- Capteur d'Accélération (Accéléromètre) :

Les accéléromètres piézoélectriques sont en train de devenir le capteur de vibration absolu le plus largement utilisé pour la surveillance. Ils ont les propriétés suivantes :

- Peut être utilisé sur une très large gamme de fréquences.
- Excellente linéarité sur une plage dynamique très large (typiquement 140 dB).
- Le signal d'accélération peut être intégré électroniquement pour donner le déplacement et la vitesse.
- Aucun élément mobile, donc extrêmement durable.

Il existe deux types de surveillance des vibrations, en ligne, où les capteurs sont installés en permanence sur la machine, et hors ligne, où la mesure est effectuée manuellement avec des capteurs.

La chaîne de mesure des vibrations doit remplir les fonctions suivantes :

- Convertir les vibrations mécaniques en signaux électriques.
- Conditionnez le signal de sortie du capteur pour qu'il soit utilisable et transportable.
- Intégrer les signaux
- plage de fréquence limitée
- Amplifier à nouveau le signal obtenu
- Signaux numérisés.

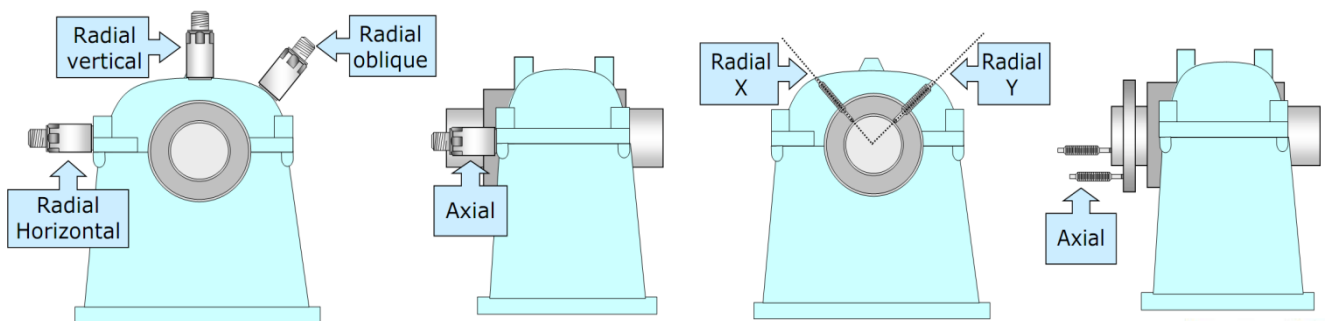


Figure I.18: Points de mesure des vibrations (absolu et relative) , [12].

I.8.3.1 Définition D'une Vibration

Un système mécanique est dit en vibration lorsqu'il est animé d'un mouvement de va-et-vient autour d'une position moyenne, dite position d'équilibre. Une vibration est généralement traduite par :

- Un déplacement : la position de la masse varie de part et d'autre du point d'équilibre.
- Une vitesse : variation du déplacement par rapport au temps.
- Une accélération : variation de la vitesse par rapport au temps, [12].

I.8.3.2 Caractéristiques D'une Vibration

a. L'Amplitude

On appelle amplitude d'un mouvement vibratoire la valeur de ses écarts par rapport à sa position d'équilibre. De cette définition générale, la complexité d'un signal vibratoire réel conduit à définir plusieurs grandeurs d'amplitude :

- **L'Amplitude crête (Ac)** : Elle représente l'amplitude maximale du signal par rapport à sa valeur d'équilibre.
- **L'Amplitude crête à crête (Acc)** : Elle représente l'écart entre les amplitudes extrêmes du signal pour un temps d'observation donné. Dans le cas d'une vibration sinusoïdale, elle est parfois appelée amplitude double. Il est noté que :

$$A_{cc} = 2A_c$$

- **L'Amplitude efficace (A_{eff}) ou RMS (Root Mean Square)** : Elle indique l'énergie donnée par le mouvement vibratoire. Pour une vibration périodique, la valeur efficace est la moyenne quadratique des valeurs efficaces de chacune des vibrations harmoniques la constituant.

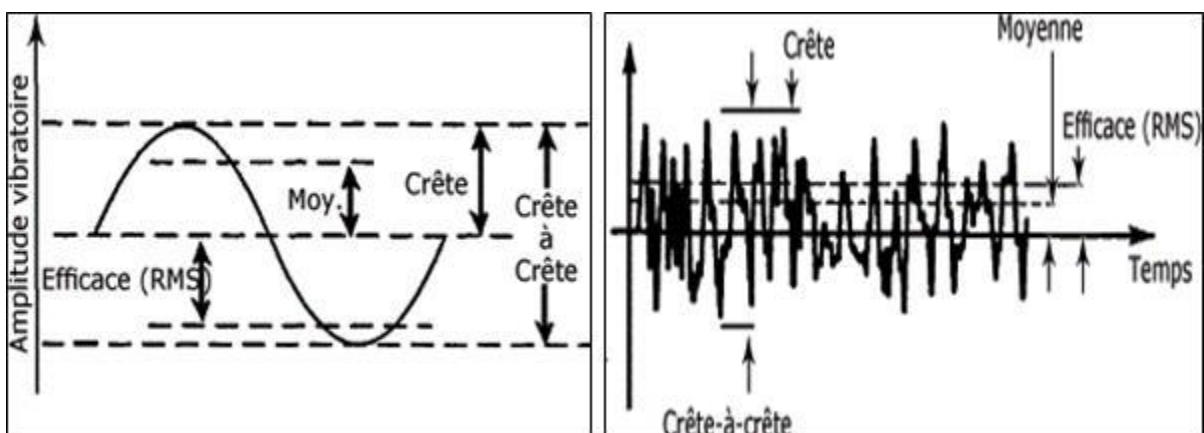


Figure I.19. Caractéristiques d'une vibration Sinusoïdale ou quelconque, [13]

b. La fréquence

La fréquence représente la cadence de répétition d'un phénomène ou le nombre de fois qu'il se reproduit en un temps donné. Lorsque l'unité de temps choisi est la seconde, la fréquence s'exprime en Hertz (Hz). Une vibration qui se produira 50 fois/seconde aura donc une fréquence de 50 Hz. La fréquence f est l'inverse de la période T qui est la durée d'un cycle.

$$f = \frac{1}{T}$$

I.8.3.3 Surveillance des vibrations :

On distingue fondamentalement deux types de surveillances des vibrations :

- Surveillance On-line.
- Surveillance Off-line.

La surveillance peut-on outre être continue ou périodique. Les systèmes on-line désignent les systèmes d'acquisition (capteurs, câbles, électronique, logiciel, ...) installés à demeure sur les machines. Ils permettent la surveillance en continu des paramètres et donc la réalisation de la fonction sécurité.



Figure I.20: Racks de surveillance des vibrations On-line

Les capteurs sont installés à demeure sur les machines et connectés à un système de surveillance comme Rack 3500 Bently Nevada, VM600 Vibromètre,

Les systèmes Off-line, les mesures sont relevées à intervalles réguliers par un opérateur équipé d'un collecteur de données comme VT60, OROS, ADRE

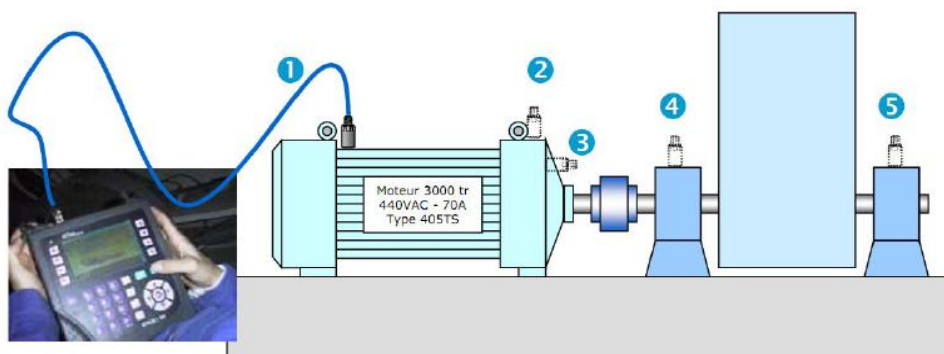


Figure I.21: Surveillance des vibrations Off-line, [14]

Les systèmes Off-line, désignent les systèmes d'acquisition (capteurs, câbles, électronique, logiciel, ...) portable permettant les relevés périodiques sur les machines est donc la réalisation de la fonction



Figure I.22: Outils de collection des données Off-line

I.9. Conclusion

Dans ce premier chapitre nous avons présenté une brève description de la turbine à gaz et son principe de fonctionnement y compris les éléments principaux constituant la turbine à gaz. Ensuite Les différents défauts que l'on peut trouver et leurs conséquences sur la turbine à gaz. Pour éviter toutes ces conséquences, nous avons vu les techniques de détection des défauts installées actuellement dans l'industrie. Notamment le défaut de vibrations qui est le plus connue et le plus largement propagé au monde des turbines à gaz car ses dégâts les plus défavorable. Donc la détection du défaut de vibrations à temps est un grand défi pour mettre les équipements de la turbine à gaz en toute sécurité et garantir une meilleure fiabilité des moyens de productions en terme de disponibilité.

L'intelligence artificielle (Techniques d'apprentissages) est un sujet d'actualité il donne à l'ordinateur la capacité « d'apprendre » à partir des données, c'est-à-dire d'améliorer leurs performances à résoudre des tâches sans être explicitement programmés pour chacune. L'efficacité et ses utilisations (L'intelligence artificielle) feront l'objet du second chapitre.

CHAPITRE II

INTRODUCTION À L'APPRENTISSAGE AUTOMATIQUE

II.1. Introduction :

L'apprentissage automatique est une application de l'intelligence artificielle (IA) qui permet aux systèmes d'apprendre et de s'améliorer automatiquement en fonction de l'expérience elle-même, sans être explicitement programmés. L'apprentissage automatique se concentre sur le développement de programmes informatiques capables d'accéder aux données et de les utiliser pour apprendre par eux-mêmes. Ce processus d'apprentissage commence par l'observation des données, telles que, une expérience de première main ou des illustrations, dans le but de trouver des modèles dans ces données et de prendre de meilleures décisions à l'avenir sur la base des exemples que nous fournissons. Notre objectif principal est de permettre à l'ordinateur d'apprendre automatiquement, sans intervention ou assistance humaine, et d'ajuster ses opérations en conséquence.

Dans le domaine de l'apprentissage automatique, il existe deux principaux types d'apprentissage: supervisé et non supervisé. La principale différence entre les deux types est que l'apprentissage supervisé est basé sur une vérité. En d'autres termes, nous avons une connaissance préalable de ce que devrait être la valeur de sortie de notre échantillon. Ainsi, le but de l'apprentissage supervisé est d'apprendre une fonction qui, étant donné un échantillon de données et le résultat souhaité, se rapproche le plus de la relation entre les entrées et les sorties qui peuvent être observées dans les données. En revanche, l'apprentissage non supervisé n'a pas de résultats étiquetés. Par conséquent, son objectif est de déduire la structure naturelle présente dans un ensemble de points de données.

II.2 Intelligence Artificielle:

L'intelligence artificielle (IA) est un domaine de l'informatique qui met l'accent sur la création de machines intelligentes qui fonctionnent et réagissent comme des humains, [15].

- L'IA est une science interdisciplinaire aux approches multiples.
- L'IA est devenue un élément essentiel de l'industrie technologique.

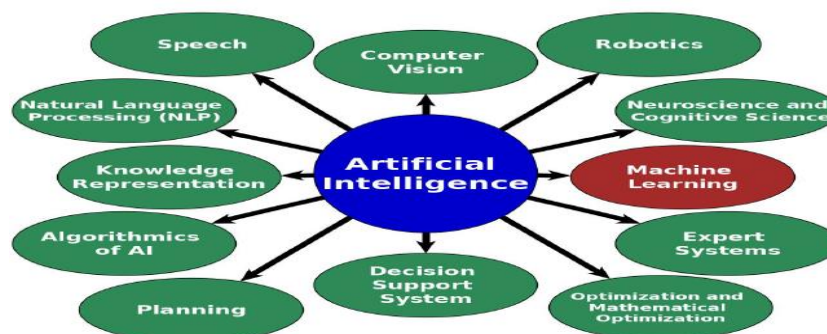


Figure II.1 : Sous-domaines de IA, [15].

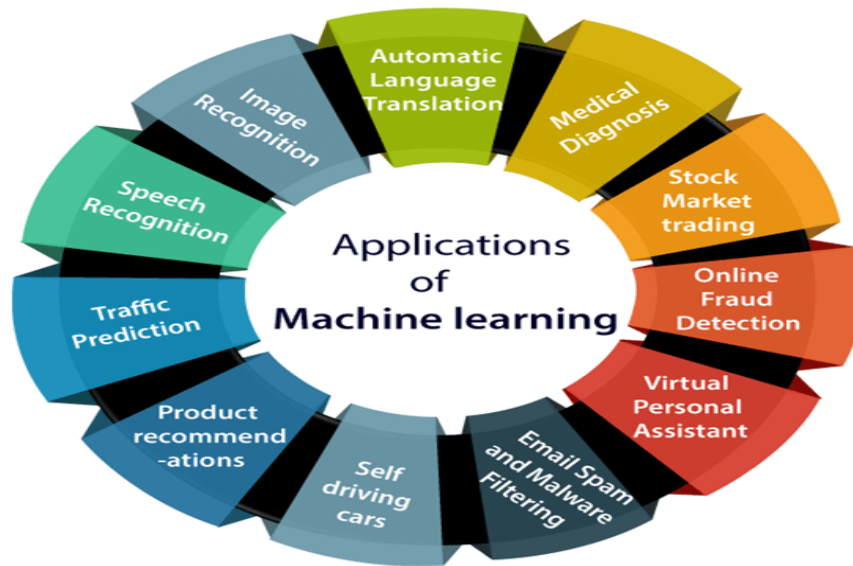


Figure II.2 : Applications des IA, [15].

II.2.1 Historique:

Cela peut sembler être une technologie relativement nouvelle, mais ce n'est pas vraiment le cas. Le premier ouvrage lié au ML a été publié en 1950, il y a 66 ans.

Aux débuts de l'IA, la recherche en ML utilisait principalement des données symboliques et la conception d'algorithmes était basée sur la logique; [16] [17]. À peu près à la même époque, Frank Rosenblatt propose le perceptron, une approche statistique basée sur la minimisation empirique du risque, [18]. Cependant, cette approche est restée méconnue et sous-développée pendant des décennies. Le véritable développement de l'apprentissage statistique a eu lieu après 1986 lorsque David Rumelhart et James McClelland ont proposé l'algorithme de rétropropagation non linéaire, [19]. Les chercheurs en IA, en reconnaissance de formes et en statistiques se sont intéressés à cette approche, et elle est maintenant largement utilisée dans les réseaux de neurones d'apprentissage en profondeur. Pour les curieux, voici une liste chronologique des événements les plus importants qui ont eu lieu dans la région depuis le "point de départ" de 1950. Pas un référent. 1950 - Alan Turing crée le "test de Turing". Ce test a déterminé si un ordinateur a une véritable intelligence. 1952 - Arthur Samuel crée le premier didacticiel informatique. Des dames étaient jouées et l'ordinateur était capable d'améliorer le jeu au fur et à mesure qu'il jouait en découvrant quelle main gagnerait la partie.

1957 - Frank Rosenblatt conçoit le premier réseau neuronal informatique (perceptron) pour simuler les processus de pensée du cerveau humain.

1967 - L'algorithme du plus proche voisin est créé, permettant aux ordinateurs de commencer la reconnaissance de formes de base.

1979 - Un étudiant de l'Université de Stanford invente le "Stanford Cart", qui peut escalader indépendamment des obstacles à l'intérieur. 1981 - Gerard DeJong introduit le concept d'apprentissage basé sur les explications (EBL). Dans ce concept, un ordinateur analyse les données d'entraînement et crée des règles générales qui peuvent être suivies en supprimant les données sans importance.

1985 - Terry Sejnowski invente NetTalk, qui apprend à prononcer des mots comme un bébé.

1986 - David Rumelhart et James McClelland proposent un algorithme de rétropropagation non linéaire.

Années 1990 - Les efforts d'apprentissage automatique sont passés d'une approche basée sur les connaissances à une approche basée sur les données. Les scientifiques commencent à développer des programmes informatiques pour analyser de grandes quantités de données (big data) et tirer des conclusions (ou "apprendre" des résultats). 1997 - Deep Blue d'IBM bat le champion du monde d'échecs. 2006 - Jeffrey Hinton a inventé le terme "apprentissage en profondeur" pour décrire de nouveaux algorithmes qui permettent aux ordinateurs de "voir" et de distinguer les objets et le texte dans les images et les vidéos.

2011 - Google Brain est développé, ses réseaux de neurones profonds apprennent à trouver et à classer des objets comme les chats.

2012 - Le X Lab de Google développe un algorithme d'apprentissage automatique capable de rechercher automatiquement des vidéos YouTube et d'identifier les vidéos contenant des chats.

2014 – Facebook développe DeepFace, un algorithme logiciel capable de reconnaître ou de vérifier les personnes sur les photos ainsi que les humains. 2015 – Amazon lance sa propre plateforme d'apprentissage automatique.

II.2.2 L'apprentissage automatique (Machine learning)

L'apprentissage automatique ou Le Machine Learning est un domaine scientifique, et plus particulièrement une sous-catégorie de l'intelligence artificielle. Elle consiste à laisser des algorithmes découvrir des « patterns, à savoir des motifs récurrents, dans les ensembles de données. Ces données peuvent être des chiffres, des mots, des images, des statistiques

Tout ce qui peut être stocké numériquement peut servir de données pour le Machine Learning. En décelant les patterns dans ces données, les algorithmes apprennent et améliorent leurs performances dans l'exécution d'une tâche spécifique.

Pour résumer, les algorithmes de Machine Learning apprennent de manière autonome à effectuer une tâche ou à réaliser des prédictions à partir de données et améliorent leurs performances au fil du temps. Une fois entraîné, l'algorithme pourra retrouver les patterns dans de nouvelles données, [20].

II.2.3 Pourquoi utiliser l'apprentissage automatique ?

Approche traditionnelle. À moins que le problème ne soit mineur, le programme peut être une longue liste de règles complexes très difficiles à maintenir. Mais c'est une approche d'apprentissage automatique. Ce programme est beaucoup plus court, plus facile à maintenir et probablement plus précis.

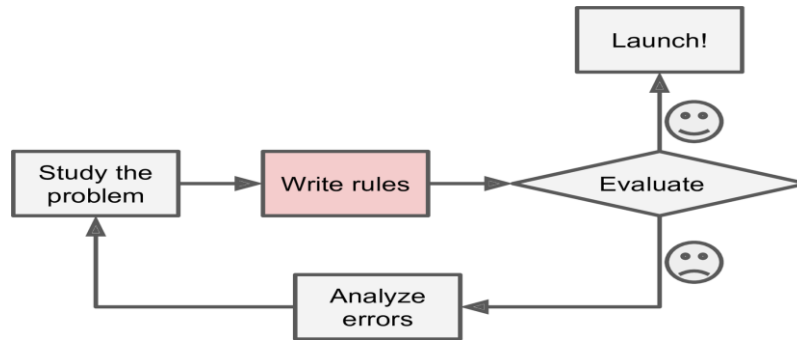


Figure II.3: L'approche traditionnelle

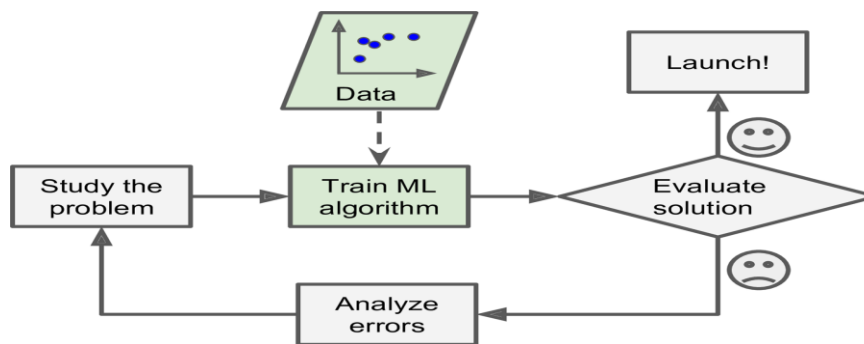


Figure II.4: Approche d'apprentissage automatique, [21].

L'apprentissage automatique aide les gens à apprendre [21] (Figure II.4): Vous pouvez revoir vos algorithmes ML pour voir ce qu'ils ont appris (bien que cela puisse être difficile pour certains algorithmes). Par exemple, une fois qu'un filtre anti-spam a été formé sur suffisamment de spam, il peut facilement vérifier une liste de mots et de combinaisons de mots qui, selon lui, prédisent le mieux le spam. Dans certains cas, cela révèle des relations inattendues et de nouvelles tendances, conduisant à une meilleure compréhension de la question.

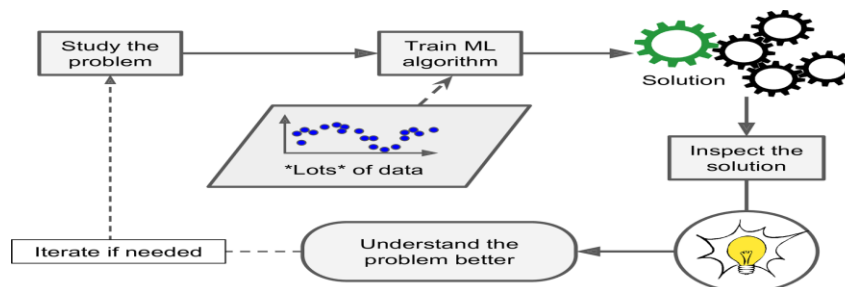


Figure II.5: L'apprentissage automatique peut aider les humains à apprendre, [21].

II.3. Types d'apprentissage automatique

Selon l'objectif à atteindre en utilisant ML, il peut être classé en trois grands types comme suit :

II.3.1. Apprentissage supervisé

L'apprentissage supervisé tire son nom du fait que le scientifique des données agit comme un guide, enseignant à l'algorithme les conclusions à tirer. Ceci est similaire à la façon dont les élèves apprennent l'arithmétique de base de leurs professeurs. Ce type d'apprentissage nécessite des données étiquetées avec l'algorithme produisant la bonne réponse attendue. Pour les problèmes de classification et de régression, l'apprentissage supervisé s'est avéré précis et rapide, [22].

- **Classification** : implique la prévision de valeurs de sortie catégorielles, où les données peuvent être classées dans des « catégories » spécifiques. La classification a différents cas d'utilisation, tels que : déterminer la météo, si un e-mail est un spam ou le type d'animal après avoir été formé sur un ensemble de données d'images qui étiquettent correctement les espèces et certaines caractéristiques d'identification, [22].

- **Régression** : il s'agit d'un problème qui nécessite la prédiction de valeurs de réponse continues telles que les actions et les prix des logements, [22]. Ainsi, cela fonctionne en modélisant la relation et les dépendances entre la sortie de prédiction cible et les caractéristiques d'entrée afin qu'il puisse prédire la valeur de sortie pour les nouvelles données en fonction de la relation qu'il a apprise des ensembles de données précédents.

II.3.1.1 Les étapes de l'apprentissage supervisé :

a. Collecter les données nécessaires :

La qualité et la quantité de données affectent directement l'efficacité des modèles génératifs. Afin de développer sa capacité à accumuler des connaissances et à prendre des décisions autonomes, les machines ont besoin de consommer une grande quantité d'informations : plus elles sont nombreuses et fiables, plus les résultats obtenus sont précis et mieux elles peuvent s'adapter aux besoins de l'entreprise. Par conséquent, les données doivent être collectées en fonction des objectifs définis à l'étape précédente. Collecter à partir de plusieurs sources de données ? Intégrez différentes bases de données en les fusionnant, [23].

b. Préparer les données :

Un modèle d'apprentissage réussi repose avant tout sur des données de qualité : il est donc nécessaire de prétraiter les données collectées afin d'en exploiter tout le potentiel. Données mal annotées, données indisponibles, informations dupliquées, incohérentes ou redondantes... L'intégration

des données peut poser de nombreux problèmes au sein d'un entrepôt de données. Ainsi, la troisième étape vise à nettoyer et standardiser (rendre comparable) les données brutes, voire à les améliorer en utilisant d'autres sources. Cible ? Rendre ces données cohérentes et disponibles pour les algorithmes. Si vous manipulez des données confidentielles, à ce stade, vous devriez envisager de les anonymiser ou de les pseudonymiser pour garantir la conformité, [23].

c. Entraînement et évaluation de modèles

De toutes les phases d'apprentissage automatique, les tests de formation restent la phase la plus distinctive de l'apprentissage automatique. Avec l'entrée de données, les modèles sont formés au fil du temps, améliorant progressivement leur capacité à répondre à des situations spécifiques, à résoudre des problèmes complexes et à accomplir des tâches. Pour cette phase d'apprentissage, nous vous recommandons d'utiliser des données d'apprentissage (également appelées "training set"). Toutes les informations collectées sont souvent trop lourdes et gourmandes en ressources.

Dans ce cas, il suffit de sélectionner une partie du jeu de données (échantillonnage) pour entraîner le modèle plus efficacement et améliorer ses prédictions. Essayez de choisir un échantillon représentatif de vos données. Sinon, vous courez le risque d'introduire un biais.

II.3.1.2 Les techniques d'apprentissage supervisé

La machine Learning, ou apprentissage automatique, est un élément principal quand il s'agit d'intelligence artificielle. L'apprentissage automatique constitue une grande avancée dès lors que vous voulez créer une intelligence artificielle ou tentez simplement d'obtenir un aperçu de toutes les données que vous avez collectées, [24]. Nous allons exposer dix (10) principaux algorithmes d'apprentissage supervisé.

1. Régression linéaire

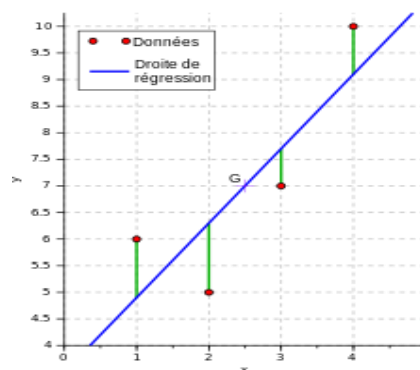


Figure II.6 : Exemple de régression linéaire simple, [24].

La régression linéaire est une des algorithmes d'apprentissage supervisé les plus populaires. En outre, il est simple et mieux compris dans les domaines des statistiques et de l'apprentissage automatique.

La régression linéaire est un type de base d'analyse prédictive. Le concept général de régression consiste à considérer deux questions:

Un ensemble de variables prédictives peut-il prédire la variable de résultat ? Quelles variables sont les plus importantes et ont le plus grand impact sur la variable de résultat ? Ces estimations de régression sont utilisées pour décrire la relation entre la variable dépendante et une ou plusieurs variables indépendantes. La forme la plus simple de l'équation de régression avec une variable dépendante et une variable indépendante est définie par la formule $y = c + b * x$. où y = variable dépendante estimée, c = constante, b = coefficient de régression, x = variable indépendante. On parle ici de régression linéaire simple. Pour la régression multiple, on écrit $y = c + b * x_1 + \dots + n * x_n$. x_1 à x_n sont les variables indépendantes et b à n sont les coefficients de régression pour chacune des variables, [24].

2. SVM (machine à vecteur de support) :

Les SVM sont une famille d'algorithmes d'apprentissage automatique qui résolvent les problèmes de classification, de régression ou de détection d'anomalies. Ils sont connus pour leurs solides garanties théoriques, leur grande flexibilité et leur facilité d'utilisation sans trop de connaissances en data mining. SVM a été développé dans les années 1990. Le principe est simple, comme le montre le schéma ci-dessous.

L'objectif est de diviser les données en classes en utilisant des frontières aussi "simples" que possible, en maximisant la distance entre les différents groupes de données et les frontières qui les séparent. Cette distance est également appelée « marge » et SVM est appelé « séparateur à large marge », les « vecteurs de support » étant les données les plus proches de la frontière.

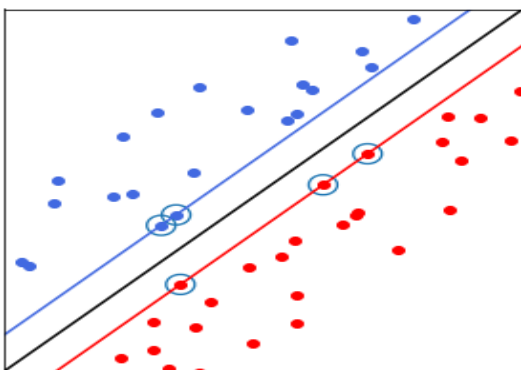


Figure II.7: Data analytic post,[24].

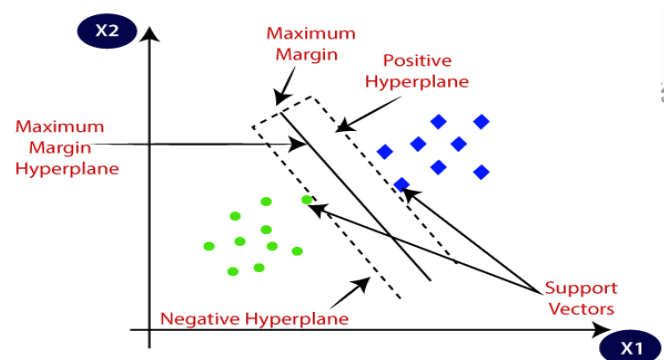


Figure II.8: Classification of data by support,[24].

Dans cet espace 2D, la "frontière" est la ligne noire, le "vecteur de support" est le point encerclé (le plus proche de la frontière) et la "marge" est la distance entre la frontière et les lignes bleues et rouges. Ce concept de limites suppose que les données sont linéairement séparables, ce qui est rarement le cas en pratique. Pour résoudre ce problème, les SVM ont souvent recours à l'utilisation d'un « noyau ». À l'aide de ces fonctions mathématiques, nous pouvons séparer les données en les projetant sur un espace de caractéristiques (espace vectoriel de grande dimension, voir figure ci-dessous). La technique de maximisation des marges assure un modèle plus robuste au bruit et un modèle plus généralisable.

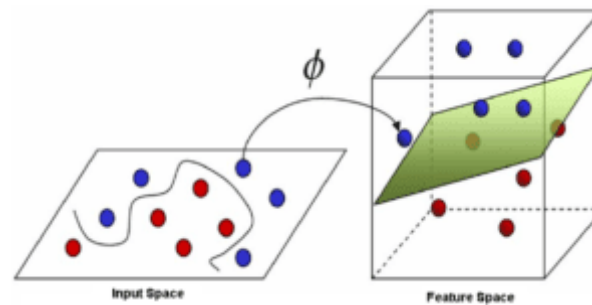


Figure II.9: Noyau,[24].

Le SVM vous permet de projeter des données dans un espace de grande dimension via les fonctions du noyau et de séparer linéairement les données.

Les SVM sont utilisées dans une variété d'applications (bioinformatique, recherche d'information, vision par ordinateur, finance, etc.), d'autant plus qu'il n'est pas nécessaire de comprendre leur fonctionnement, contrairement aux réseaux de neurones. Il existe des ensembles d'hyperparamètres standards pour la classification, la régression ou la détection d'anomalies qui fonctionnent dans la plupart des cas. C'est l'un de leurs principaux avantages. De plus, il existe peu d'hyperparamètres tels que :

Vous êtes limité au choix de la méthode de régularisation (type Lasso, ou régularisation RKHS*, qui est une méthode propre à SVM) et au choix du noyau (noyau polynomial, Sobolev, RBF**, etc.). En ce qui concerne les algorithmes SVM, ils sont appelés régression de crête du noyau pour la régression et SVM à une classe pour la détection d'anomalies. Enfin, nos données montrent que les SVM fonctionnent généralement aussi bien ou mieux que les réseaux de neurones et les modèles de mélange gaussien, à l'exception de certains cas saillants tels que la classification d'images. Il a également été montré qu'en utilisant le noyau RBF, le SVM devient un « approximateur universel » [24]. H. Avec suffisamment de données, l'algorithme peut toujours trouver la meilleure frontière pour séparer les deux classes (en supposant que la frontière existe).

3. Arbre de décision

Les arbres de classification et de régression, également appelés CART (arbres de classification et de régression), sont une forme simple d'arbre de décision. Seuls les algorithmes et les structures de données sont utilisés dans cette structure. Ces arbres se composent de seulement deux composants:

- Nœud de branche. Représente une variable d'entrée unique et fournit un point de partage unique pour la variable.
- Un nœud feuille représentant deux variables de sortie, [25].

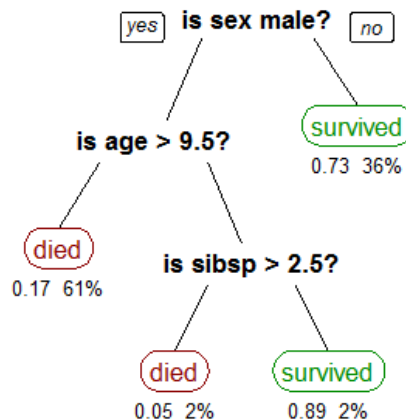


Figure II.10:Un arbre de décision,[25].

Lorsque l'algorithme est exécuté par une machine, les prédictions sont faites en suivant les subdivisions des nœuds de branche jusqu'à ce qu'ils atteignent les nœuds feuilles. Ce nœud feuille est la prédiction ou la sortie de la valeur de classe.

Les arbres de classification et de régression sont faciles à apprendre et à utiliser, et précis pour une grande variété de problèmes. Celles-ci sont particulièrement rapides à mettre en œuvre, car les données ne nécessitent aucune préparation particulière, [25].

4. K- plus proches voisins :

Les algorithmes K-NN (abréviation de k-plus proches voisins) utilisent l'intégralité de l'ensemble de données comme apprentissage au lieu de le diviser en ensembles d'apprentissage et de test.

Lorsque les résultats d'une nouvelle instance de données sont nécessaires, l'algorithme KNN parcourt l'ensemble de données entier pour trouver les k instances les plus proches de la nouvelle instance ou les plus similaires au nouvel ensemble de données, et calcule la moyenne des résultats à renvoyer, [24]. , ou dans le cas de problèmes de classification, la classe à laquelle appartient cette instance. Vous spécifiez vous-même la valeur de k.

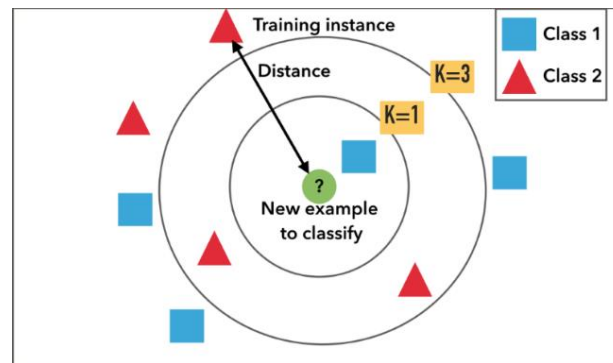


Figure II.11: K-plus proches voisins (apprentissage paresseux), [24].

5. Forêts aléatoires :

Le principe de la forêt aléatoire est simple, il suffit de faire pousser des arbres. Mais bien sûr, créer 100 arbres avec les mêmes données signifie générer le même arbre 100 fois. Nous modifions donc légèrement les données pour chacun des 100 arbres.

Une fois les échantillons générés, nous pouvons construire l'arbre. Pour ce faire, sélectionnez au hasard des colonnes spécifiques à partir des données pour éviter d'utiliser toutes les colonnes à la fois et entraînez-vous sur chaque échantillon bootstrap. Vous avez donc 100 arbres et ils sont tous légèrement différents.

Une fois la formation terminée, les 100 règles de décision de l'arbre sont enregistrées pour une utilisation en mode prédictif. Cela revient à exécuter de nouvelles observations sur 100 arbres, chaque arbre définissant une classe d'appartenance. Un vote majoritaire est utilisé pour déterminer la classe prédite.

6. Réseau neurones artificiel :

Les réseaux de neurones sont des modèles complexes il peut être utilisé dans l'apprentissage supervisé pour la régression et la classification. Il est de plus en plus utilisé dans de nombreux domaines car il permet la modélisation du phénomène complexe. Par exemple, vous pouvez les trouver dans le traitement d'image. Reconnaissance de formes pour le traitement linguistique et l'approximation Fonctions complexes qui prennent beaucoup de temps à calculer, [26].

7. Adaboost

La méthode Adaboost (Adaptive Boosting) permet de combiner plusieurs règles Créez-en simplement un plus efficace. L'algorithme Adaboost a été proposé Dans Freund et Schapire (1996). L'idée de base du Boosting est de combiner 71 "règles" simples pour créer un ensemble dont la performance pour chaque élément est mieux que dans l'ensemble. IL existe plusieurs variantes de la méthode Boosting, dont la méthode AdaBoost. M 1 est conçue pour Classification. Cette méthode peut

utiliser n'importe quel algorithme d'apprentissage Classification. Cependant, l'algorithme d'apprentissage doit pouvoir gère les instances pondérées (le poids doit être un nombre positif). La présence des poids d'instances change la manière de calculer l'erreur par le classificateur qui correspond à la somme des poids des instances mal classifiées divisée par le poids total de toutes les instances, au lieu de la fraction des instances mal classifiées, [27].

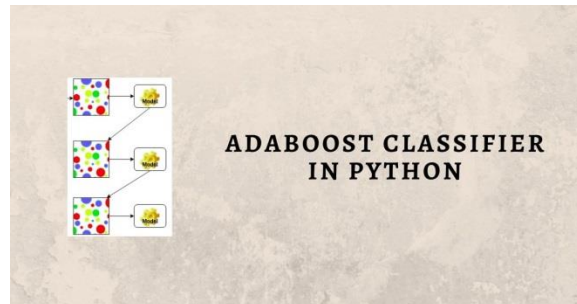


Figure II.12 : Algorithme Adaboost,[27].

8. XGBoost

XGBoost est un acronyme pour Extrême Gradient Boosting, un algorithme d'apprentissage automatique populaire qui appartient à la famille des méthodes de gradient boosting. il a été développé Proposé par Tian qi Chen en 2014, il a gagné en popularité en raison de son efficacité et de sa polyvalence dans la résolution d'un large éventail de problèmes d'apprentissage automatique. XGBoost est une méthode d'apprentissage d'ensemble qui combine les prédictions de plusieurs apprenants faibles (généralement des arbres de décision) pour créer un modèle prédictif puissant. Il fonctionne en ajoutant de manière itérative de nouveaux arbres de décision pour corriger les erreurs commises par les arbres précédents. Ce processus itératif minimise une fonction de perte qui mesure la différence entre les valeurs prédites et réelles de la variable cible, [27].

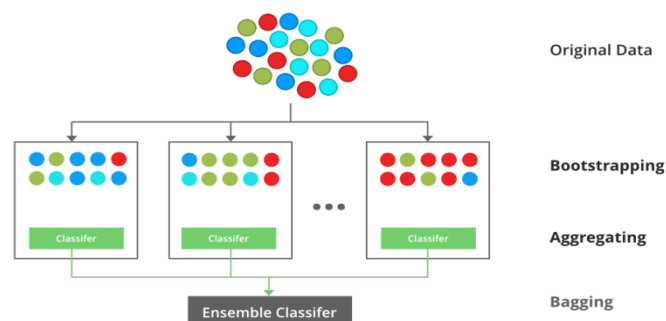


Figure II.13 : Algorithme XGboost,[27].

9. Gradient Boosting :

Est un puissant algorithme de boosting qui combine plusieurs apprenants faibles en apprenants forts, dans lequel chaque nouveau modèle est formé pour minimiser la fonction de perte telle que l'erreur quadratique moyenne ou l'entropie croisée du modèle précédent en utilisant la descente de gradient. A chaque itération, l'algorithme calcule le gradient de la fonction de perte par rapport aux prédictions de l'ensemble courant puis forme un nouveau modèle faible pour minimiser ce gradient. Les prédictions du nouveau modèle sont ensuite ajoutées à l'ensemble, et le processus est répété jusqu'à ce qu'un critère d'arrêt soit satisfait.

Contrairement à AdaBoost, les pondérations des instances de formation ne sont pas modifiées, mais chaque prédicteur est formé en utilisant les erreurs résiduelles du prédécesseur comme étiquettes. Il existe une technique appelée Gradient Boosted Trees dont l'apprenant de base est CART (Classification and Regression Trees). Le diagramme ci-dessous explique comment les arbres à gradient renforcé sont entraînés pour les problèmes de régression.

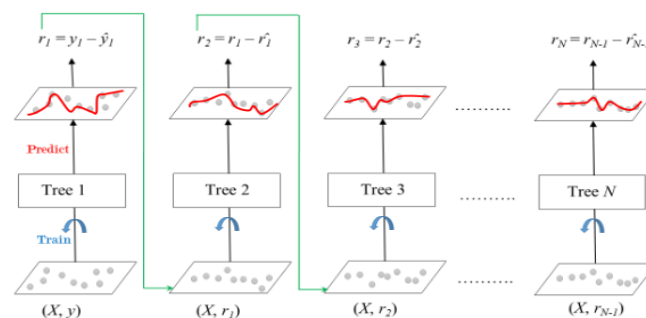


Figure II.14 : Gradient boosted trees pour la régression,[28].

L'ensemble est constitué de M arbres. Tree1 est formé à l'aide de la matrice de caractéristiques X et des étiquettes y . Les prédictions étiquetées \hat{y}_1 sont utilisées pour déterminer les erreurs résiduelles de l'ensemble d'apprentissage r_1 . Tree2 est ensuite formé en utilisant la matrice de caractéristiques X et les erreurs résiduelles r_1 de Tree1 comme étiquettes. Les résultats prédits \hat{r}_1 sont ensuite utilisés pour déterminer le résidu r_2 . Le processus est répété jusqu'à ce que tous les arbres M formant l'ensemble soient entraînés. Il existe un paramètre important utilisé dans cette technique connue sous le nom de rétrécissement. La réduction fait référence au fait que la prédiction de chaque arbre de l'ensemble est réduite après avoir été multipliée par le taux d'apprentissage (η) qui varie entre 0 et 1. Il existe un compromis entre η et le nombre d'estimateurs, ce qui diminue l'apprentissage. Le taux doit être compensé par des estimateurs croissants afin d'atteindre certaines performances du modèle. Étant donné que tous les arbres sont formés maintenant, des prédictions peuvent être faites. Chaque arbre prédit une étiquette et la prédiction finale est donnée par la formule, [28].

10. Naive Bayes :

La méthode de classification naïve bayésienne est un algorithme d'apprentissage supervisé (supervised machine learning) qui permet de classifier un ensemble d'observations selon des règles déterminées par l'algorithme lui-même. Cet outil de classification doit dans un premier temps être entraîné sur un jeu de données d'apprentissage qui montre la classe attendue en fonction des entrées. Pendant la phase d'apprentissage, l'algorithme élabore ses règles de classification sur ce jeu de données, pour les appliquer dans un second temps à la classification d'un jeu de données de prédiction. Le classificateur bayésien naïf implique que les classes du jeu de données d'apprentissage soit connu et fournit, d'où le caractère supervisé de l'outil.

Historiquement, la classification naïve bayésienne fut utilisée pour la classification de documents et l'élaboration de filtres anti-spam. Aujourd'hui, c'est un algorithme renommé dont les applications peuvent être rencontrées dans de nombreux domaines. Parmi ces atouts les plus significatifs, on citera son apprentissage rapide qui ne nécessite pas un gros volume de données et son extrême rapidité d'exécution comparé à d'autres méthodes plus complexes. Finalement, malgré la forte hypothèse simplificatrice d'indépendance des variables (voir description ci-dessous), la classification naïve bayésienne obtient des résultats remarquables dans de nombreuses applications de la vie courante ce qui en fait un algorithme de choix parmi les outils du machine learning, [29].

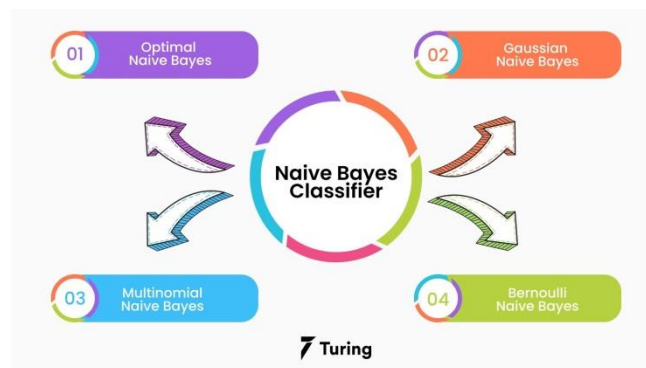


Figure II.15 : Algorithme Naive Bayes,[29].

II.3.2. Apprentissage non supervisé

L'apprentissage non supervisé est une branche de l'apprentissage automatique caractérisée par l'analyse et l'agrégation de données non étiquetées. Pour ce faire, ces algorithmes apprennent à trouver des modèles et des groupes dans les données sans intervention humaine. Mathématiquement parlant, l'apprentissage non supervisé consiste à observer plusieurs occurrences d'un vecteur X et à apprendre la distribution de probabilité $p(X)$ de ces occurrences.

II.3.2.1 Types de problèmes d'apprentissage non supervisé

On peut considérer l'apprentissage non supervisé comme étant séparé en deux catégories : le clustering et l'association.

Le clustering est une technique d'apprentissage non supervisé, qui groupe des données non-étiquetées en fonction de leurs similarités et de leurs différences. Ainsi, les points sont rassemblés dans des groupes (ou des clusters) de telle manière à ce que les points au sein d'un même groupe soient le plus similaire possible, pendant que les points dans des groupes différents soient les plus différents possible.

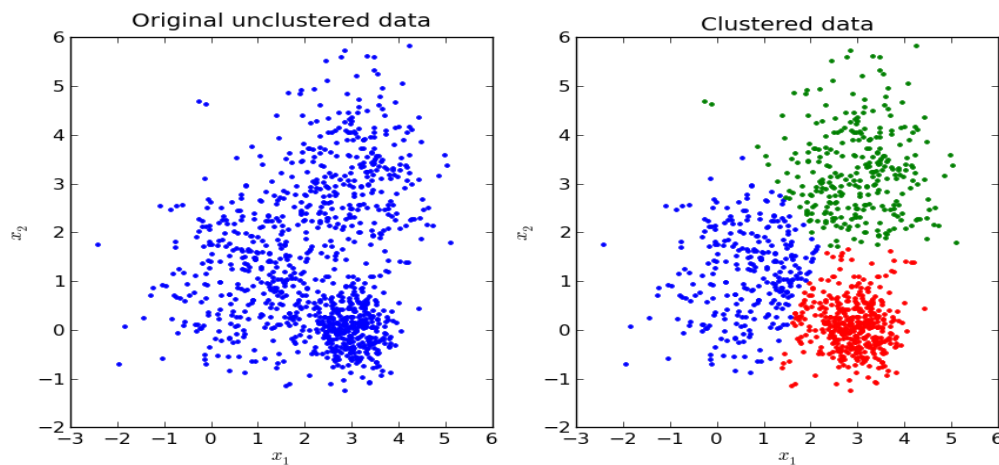


Figure II.16 : Apprentissage en profondeur clustering non supervisé,[30].

Les règles d'association, d'autre part, sont une autre forme d'apprentissage non supervisé qui identifie les relations entre les points de données. En d'autres termes, ces algorithmes trouvent des points qui apparaissent ensemble dans les données, [30].

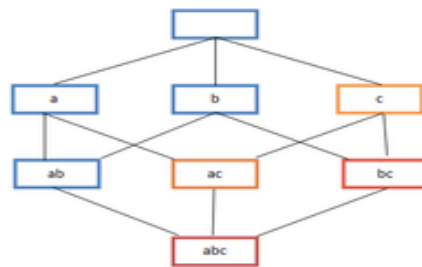


Figure II.17: Treillis d'itemset fréquents,[30].

II.3.2.2 Applications de l'apprentissage non supervisé

Malgré le fait que l'apprentissage non supervisé a tendance à être plus difficile d'utilisation, cette méthode reste néanmoins utilisée pour diverses applications.

a. Détection d'anomalies dans la finance

Les algorithmes d'apprentissage non supervisé ont la capacité d'analyser d'importantes quantités de données et d'identifier les points inhabituels au sein de ces données, [30].

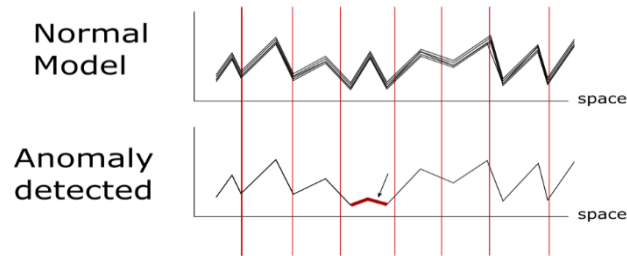


Figure II.18 : Un algorithme de détection d'anomalies,[30].

La détection d'anomalies peut également avoir d'autres applications comme pour l'identification d'erreurs humaines ou des défaillances de sécurité.

b. Clustering

Dans le domaine médical, de grandes quantités de données non-étiquetées sont disponibles. Dans ce domaine, l'apprentissage non supervisé est privilégié car l'étiquetage des données médicales peut se montrer très coûteux, en termes de temps et d'argent. Ainsi, les modèles de clustering peuvent être très utiles pour la détection, la segmentation et la classification d'image. En effet, les algorithmes de clustering peuvent recevoir une grande quantité de données et identifier des groupes ou des schémas au sein de ces informations, chose qui pourrait être très difficile pour un professionnel, [30].

c. Moteurs de recommandations et publicité personnalisées

En utilisant l'apprentissage non supervisé, un modèle peut recevoir un achat ou l'historique d'achat d'un individu et l'utiliser pour identifier des schémas et faire des prédictions. Ainsi, des compagnies peuvent utiliser ce type de modèle pour mettre au point des stratégies de vente adaptées et des publicités ciblées, en fonction de des données d'un individu.

II.3.2.3 Les techniques d'apprentissage non supervisé :

a. Méthode des K-Means

La méthode de clustering Graphical K-Means ou K-Means est l'un des algorithmes de clustering les plus couramment utilisés. Dans cette méthode, les points de données sont affectés à K groupes. Où K représente le nombre de groupes créés en fonction de la distance entre les noyaux de chaque groupe. Ce noyau, également appelé centroïde, est choisi aléatoirement en fonction des données ou spécifié par

le data scientist. Une fois que le nombre et les centres de gravité des clusters (K) sont identifiés, le modèle mappe chaque nouveau point sur le noyau et le cluster les plus proches dans le cluster approprié. La méthode la plus courante pour calculer la distance entre un point et un noyau est la distance euclidienne au carré, [31].

L'un des principaux défis de la méthode K-Means est la nécessité de spécifier le nombre de clusters avant de démarrer le modèle. Cela peut devenir compliqué, surtout si la quantité de données est très importante. Si vous choisissez une valeur élevée pour K , le modèle restitue des groupes plus petits. En revanche, lorsque K prend de petites valeurs, le modèle restitue de grands groupes.

Cette méthode est le plus souvent utilisée pour la classification de documents, la segmentation d'images et la segmentation marketing.

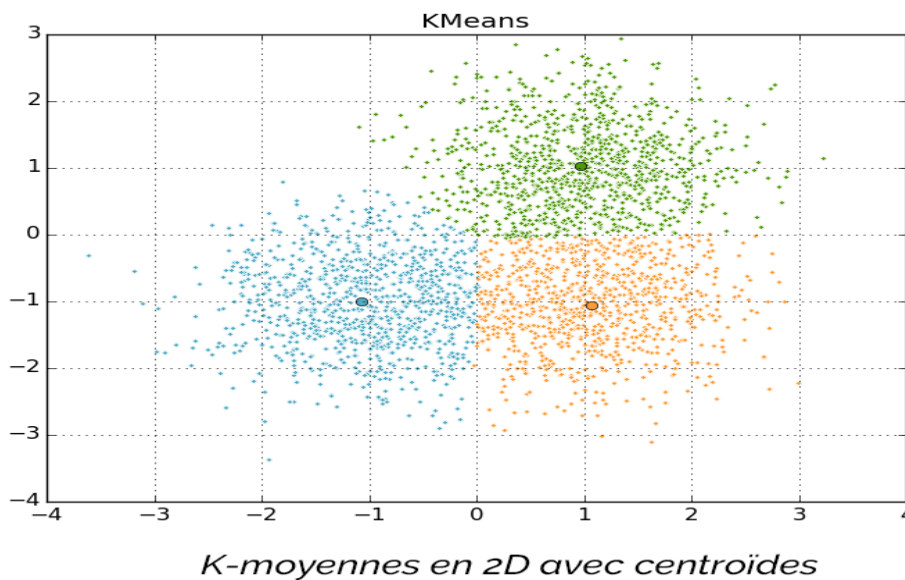


Figure II.19: La méthode des K-moyennes, ou K-means clustering.[31].

b. Algorithme à priori:

Les algorithmes a priori font partie des algorithmes d'apprentissage non supervisé qui peuvent être classés comme règles associatives. En fait, cette technique utilise une approche ascendante, identifiant les points ou collections de points les plus communs et les utilisant pour établir des règles d'association. Cet algorithme est basé sur l'idée que les sous-groupes de groupes communs sont également des groupes communs, [31].

Les algorithmes a priori sont devenus populaires pour être utilisés dans l'analyse des paniers d'achat et les recommandations musicales pour les applications populaires. En effet, grâce à cet algorithme, on peut déterminer la probabilité qu'une personne achète ou regarde le produit X sachant qu'elle a acheté ou regardé le produit Y. Par conséquent, le modèle nécessite le comportement passé d'une personne afin de pouvoir prédire son comportement futur.

II.3.3. Apprentissage par renforcement :

L'apprentissage par renforcement est le troisième type majeur d'apprentissage automatique, avec l'apprentissage supervisé et non supervisé. Il se compose de cinq éléments importants. Agents, environnements, états, actions, récompenses. L'objectif de RL (Reinforcement Learning) est de maximiser la récompense et de minimiser les risques en tirant parti de l'interaction avec l'environnement. Les algorithmes RL (appelés agents) sont régulièrement améliorés en examinant l'environnement à travers différents états possibles. L'agent détermine automatiquement le comportement idéal pour optimiser les performances. Le feedback (récompenses) permet aux agents d'améliorer leur comportement.

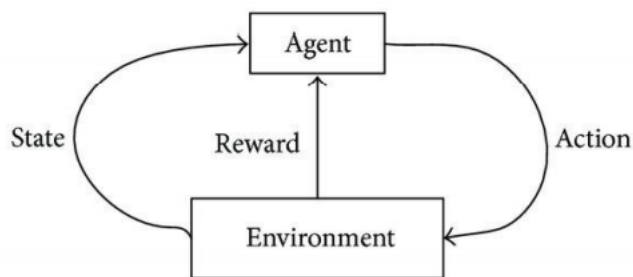


Figure II.20: Reinforcement Learning Components,[31].

L'apprentissage par renforcement profond est utilisé dans de nombreux domaines.

- La robotisation des usines et des entrepôts permettra aux machines d'apprendre à placer de nouveaux modèles de pièces sans programmation préalable.
- Étalonnage et contrôle de la qualité des systèmes industriels tels que la fabrication, la chaîne d'approvisionnement et la production d'énergie.
- la trésorerie pour optimiser le trading automatisé ou la gestion des risques de marché.
- Synthèse textuelle pour estimer la qualité globale des résumés en les extrayant d'une logique mot à mot.
- Un moteur de jeu et de recommandation pour développer des stratégies dans des environnements incertains.
- Les voitures autonomes visent à améliorer la capacité du véhicule à réagir à des événements de circulation spécifiques.

II.3.3.1 Les étapes de l'apprentissage par renforcement :

Il y a essentiellement quatre étapes à retenir dans ce domaine de l'apprentissage automatique. Notez que les agents ici peuvent faire référence à des programmes, des robots ou d'autres entités virtuelles. Voici les 4 étapes.

Surveillance de l'état environnement : Les agents surveillent l'état de l'environnement dans lequel ils se trouvent.

Veillez sélectionner une action : L'agent choisit l'action à entreprendre en fonction de l'état actuel de l'environnement (généralement interactif).

Gagner des récompenses : L'agent reçoit des récompenses en fonction des actions qu'il choisit et de l'état de l'environnement dans lequel il se trouve.

Mise à jour de la politique : Les agents utilisent les récompenses qu'ils reçoivent pour mettre à jour leurs politiques de prise de décision et choisir de meilleures actions la prochaine fois.

II.3.3.2-Les techniques d'apprentissage par renforcement:

a. Q-Learning

Peut effectuer des actions aléatoires quelle que soit la politique en vigueur. L'algorithme vise principalement à apprendre des politiques qui maximisent la récompense globale. Le Q-Learning met l'accent sur l'utilité des actions entreprises pour recevoir des récompenses. Il s'agit d'une technique fortement recommandée pour commencer l'apprentissage par renforcement.

Par exemple, lors de l'apprentissage d'un jeu, un algorithme d'apprentissage Q choisit d'abord d'explorer ou d'utiliser des données. Les agents d'IA peuvent choisir d'explorer s'ils n'ont aucune expérience. Sinon, l'exploration de données est utilisée. Après avoir pris des mesures, l'agent passe à l'étape suivante, où il prend une nouvelle décision. Au fur et à mesure que l'agent IA apprend, il progresse lentement en fonction de l'expérience qu'il acquiert.

b. Deep Q- networks

Le Deep Q- networks a été développé par DeepMind en 2015. Il a été en mesure de résoudre un large éventail de jeux Atari (dont certains à un niveau surhumain) en combinant l'apprentissage de renforcement et les réseaux de neurones profonds à l'échelle. L'algorithme a été développé par l'amélioration d'un algorithme de RL classique appelé Q-Learning avec les réseaux de neurones profonds et une *rediffusion* de l'expérience technique appelée,[31].

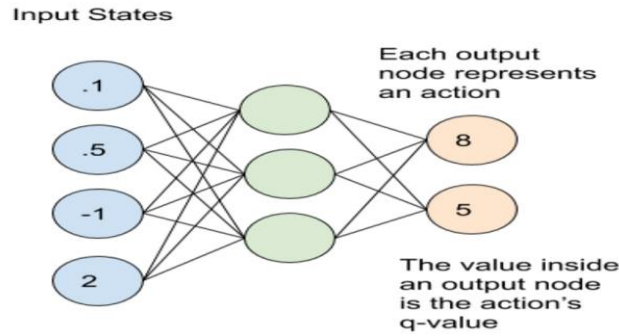


Figure II.21 : Un réseau de neurones mappant un état d'entrée à sa paire correspondante (action, valeur q) (Image de l'auteur) ,[31].

II.3.3.3. Évaluation et sélection de modèle :

Les métriques d'évaluation sont propres au type de tâche de machine learning effectuée par un modèle.

a. Courbe ROC

Supposons que vous ayez un classificateur qui classe les observations en classes. De plus, il y a aussi cette réponse avec un score de pertinence. Il y a deux cas possibles:

Soit la réponse est correcte (1), soit la réponse est incorrecte (0). Attribuez une paire (r,x) pour chaque observation. Où r vaut 0 ou 1. x est la valeur de pertinence. Je veux déterminer au-dessus de quel seuil de pertinence la réponse du classificateur est fiable. Faire varier x donne la courbe

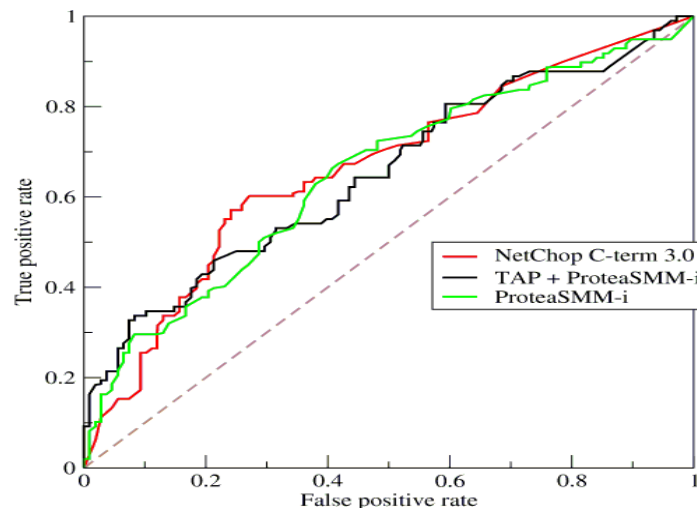


Figure II.22 : La courbe ROC de trois estimateurs d'épitope

Cette courbe est également utilisée pour comparer différents classificateurs. Plus la valeur de la courbe est élevée, plus la zone sous la courbe est grande et moins le classificateur fait d'erreurs.

En termes simples, le classificateur renvoie une mauvaise (-) ou une bonne (+) réponse. Pour créer un classificateur, nous avons toujours une base contenant la réponse attendue afin de pouvoir l'évaluer. En fonction du score x et du seuil.

b. Recall

Le rappel donne le pourcentage de résultats positifs qui ont été correctement prédits par le modèle. En d'autres termes, c'est le nombre de positifs bien prédits (vrais positifs) divisé par le nombre total de positifs (vrais positifs + faux négatifs). Mathématiquement ça ressemble à ça :

Mais à quoi sert la mémoire ?

Plus cette valeur n'est élevée, plus le modèle de machine learning maximise le nombre de vrais positifs.

Notez cependant que cela ne signifie pas que le modèle n'est pas faux. Un rappel élevé signifie que vous ne manquez probablement aucun élément positif. Cependant, il n'y a aucune information sur la qualité des prédictions négatives.

$$recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

c. Précision

La précision est très similaire à la mémoire, il est donc important de comprendre la différence.

Affiche le nombre de prédictions positives réussies.

C'est-à-dire le nombre de positifs bien prédits (vrais positifs) divisé par tous les positifs prédits (vrais positifs + faux positifs).

Exprimé mathématiquement, cela ressemble à ceci :

A quoi sert la précision ?

Plus que cette valeur est élevée, plus le modèle de machine learning minimise le nombre de faux positifs. Une grande précision signifie que la majorité des prédictions positives du modèle sont des résultats positifs bien prédits, [32].

$$precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

d. F-mesure

Bien qu'utiles, ni la précision ni le rappel ne peuvent évaluer pleinement un modèle d'apprentissage automatique. Ces deux métriques en elles-mêmes sont inutiles.

Le rappel est élevé si le modèle prédit toujours "positif".

Inversement, si le modèle ne prédit pas « positif », la précision est élevée.

Ainsi, nous obtenons des métriques qui indiquent que le modèle est efficace mais simple plutôt qu'intelligent. Heureusement, il existe des métriques qui combinent précision et rappel, [32].

Résultat F1. Le score F1 donne une bonne évaluation des performances du modèle.

Calculé comme :

$$F1 \text{ Score} = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}}$$

e. La validation croisée

La validation croisée est une méthode de test des performances des modèles d'apprentissage automatique prédictifs. Apprenez les techniques les plus couramment utilisées et comment les maîtriser.

Les modèles d'apprentissage automatique sont formés sur des données étiquetées, puis exécutés sur de nouvelles données. Cependant, il est important de garantir l'exactitude des prédictions du modèle dans un environnement de production.

Cela nécessite la validation du modèle. Le processus de validation consiste à déterminer si les résultats numériques qui quantifient les relations hypothétiques entre les variables sont des explications acceptables pour les données.

Pour évaluer les performances de votre modèle d'apprentissage automatique, vous devez le tester par rapport à de nouvelles données. En fonction des performances du modèle sur des données inconnues, vous pouvez déterminer si le modèle est sous-ajusté, surajusté ou "bien généralisé".

Une technique pour tester l'efficacité des modèles d'apprentissage automatique est la "validation croisée". Cette méthode est également une technique de "rééchantillonnage" pour évaluer des modèles même avec des données limitées. Pour effectuer "CV" (validation croisée), nous devons éliminer la partie en amont des données de l'ensemble de données d'apprentissage. Ces données ne sont pas utilisées pour la formation du modèle, mais plus tard pour le test et la validation du modèle.

La validation croisée est largement utilisée dans l'apprentissage automatique pour comparer différents modèles et choisir le plus approprié pour un problème donné. C'est plus facile à comprendre, plus facile à mettre en œuvre et moins biaisé que d'autres méthodes, [32].

f. La régularisation

Cette importante technique d'apprentissage automatique vise à limiter le "surajustement" et à contrôler les erreurs de type variance pour améliorer les performances.

Lors de l'apprentissage d'un modèle, la régularisation permet d'imposer des contraintes favorisant les modèles simples au détriment des modèles complexes. En d'autres termes, l'erreur de type variance est réduite et la généralisation de la solution est améliorée. La régularisation peut prendre plusieurs formes, en fonction de l'objectif recherché et des hypothèses faites sur le problème. Ainsi, dans la régression des moindres carrés, la régularisation euclidienne privilégie les faibles coefficients, tandis que la régularisation Lasso (utilisée lorsque le nombre de variables d'entrée est important) force l'algorithme à n'utiliser que des petits coefficients et donne la priorité à la « parcimonie » de la représentation. Certaines des données considérées sont ignorées. Pour les réseaux de neurones, les méthodes de régularisation les plus courantes sont le dropout (où les poids, les paramètres du réseau de neurones, sont aléatoirement remplacés par des 0 lors de l'apprentissage) et l'arrêt précoce (où l'apprentissage est démarré tôt au profit de modèles plus simples), ou la régularisation euclidienne précitée, [32].

De plus, la régularisation peut généralement garantir certaines propriétés théoriques de l'algorithme, garantissant ainsi de bonnes performances de l'algorithme, telles que la stabilité de la méthode du noyau et des bornes de généralisation améliorées.

g. Accuracy: L'accuracy permet de connaître la proportion de bonnes prédictions par rapport à toutes les prédictions. L'opération est simplement : Nombre de bonnes prédictions / Nombre total de prédictions.

$$Accuracy = \frac{True\ Positif + True\ Négatif}{True\ Positif + True\ Négatif + False\ Positif + False\ Négatif}$$

VI. Conclusion

Dans ce chapitre nous avons d'abord abordé les deux techniques d'apprentissages automatiques supervise et non supervise et la différence entre les deux ainsi l'apprentissage par renforcement et ses techniques et enfin on a vu l'évaluation et la sélection du modèle.

Le chapitre suivant objet d'une focalisation sur la détection des défauts de la turbine à gaz.

CHAPITRE III

RÉSULTATS ET DISCUSSIONS

III.1.Introduction

Les défauts sur les turbines à gaz provoquent des perturbations affectant leurs fonctionnements et ses qualités de service. La détection de ces défauts est l'objet de recherche depuis des décennies. Au début, la recherche a été principalement concentrée sur l'aéronautique ensuite, elle a été appliquée sur d'autres secteurs comme le nucléaire. La littérature sur le sujet s'est enrichie grâce à l'apport de différents groupes de recherches qui ont proposé de variétés d'approches de la détection de défaut basées sur l'expertise de leur propre domaine, [33]. Des améliorations de ces méthodes ont été également enrichies en réponse aux exigences de l'industrie en matière de sûreté, fiabilité et automatisation des systèmes. Les systèmes de protection surveillent en permanence l'état d'une turbine à gaz et provoquent une mise hors service, lorsque ces turbines sont le siège d'une perturbation indésirable: haute température, défaut de vibrations, mauvaise de combustion,...etc.

L'anticipation de détection des défaillances est une fonction importante en maintenance prédictive. L'analyse vibratoire est la méthode la plus largement utilisée pour la surveillance, le diagnostic fiable et pour la détection d'apparition et du développement de la plupart des défauts d'origines diverses.

La méthodologie expérimentale comprend la collecte de données de vibration, à partir des capteurs de vibrations qui sont installés sur les paliers de la turbine à gaz V94.3A de l'unité de M'sila. dans différentes conditions de fonctionnement et de scénarios de défauts. Ces données sont ensuite prétraitées en utilisant les fonctionnalités fournies par les bibliothèques de traitement de données de Python, telles que NumPy et Pandas.

Ensuite, les algorithmes d'apprentissage automatique, tels que les réseaux neuronaux, les machines à vecteurs de support ou les arbres de décision, sont utilisés pour l'entraînement et les tests des modèles de détection et de classification des défauts.

III.2 Diagnostic des défauts des vibrations en vigueur :

Toutes les machines vibrent et le spectre des fréquences de leurs vibrations a un profil particulier lorsqu'elles sont en état de "bon fonctionnement". Mais dès que les phénomènes d'usure, de fatigue, de vieillissement apparaissent, le profil de ce spectre se modifie.

Le signal vibratoire extrait d'une machine tournante est très complexe, originaire des différents organes. La transformée de Fourier est un outil mathématique permettant de transformer ces signaux complexes, de les décomposer en une multitude de composantes élémentaires sinusoïdales et de les

représenter sous forme d'un spectre « Amplitudes-fréquence » [23] [22]. La transformée de Fourier rapide FFT développée par James Cooley et John Tukey, ne requiert qu'un temps de calcul minime pour appliquer la transformée de Fourier Discrète. La FFT a été largement implémentée dans les systèmes de diagnostic des machines tournantes, pour diagnostiquer l'origine et le type de défaut.

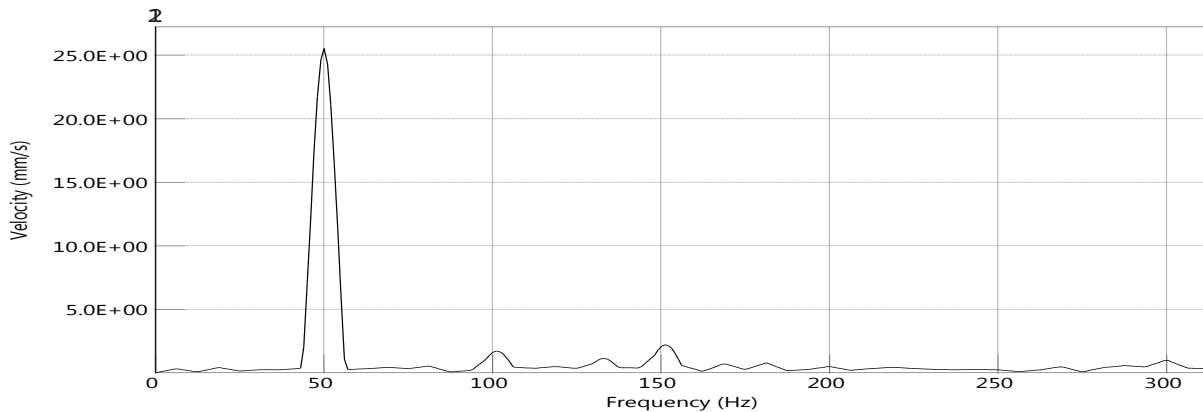


Figure III.1. Décomposition en FFT d'un signal de vibration,[33].

III.3. Critères D'évaluation

Les normes ISO de la série ISO 10816 « Vibrations mécaniques - évaluation des vibrations des machines par mesurages sur les parties tournantes » et les séries ISO 7919 « Vibrations mécaniques des machines non alternatives - mesurages sur les arbres tournants et critères d'évaluation » ont été révisées au long de la dernière décennie et offrent pour la première fois des informations sur l'interprétation des résultats d'évaluation et permettant d'établir des valeurs limites à la surveillance des vibrations. Se remettre à la partie 4 des deux normes ainsi qu'à leur introduction générale comprenant la partie I pour toute information pertinente aux turbines à gaz. Au lieu d'indiquer les limites de la plage de vibration, les deux séries désignent les plages en question par des zones marquées avec les lettres A, B, C et D. Des évaluations ont été effectuées sur chaque zone et s'appliquent régulièrement dans toutes les parties des séries. Les zones sont définies de la manière suivante :

- Zone A : la vibration de machines récemment mises en marche est normalement incluse dans cette zone.
- Zone B : les machines produisant des vibrations incluses dans cette zone sont jugées convenables pour des opérations prolongées sans restrictions.
- Zone C : les machines produisant des vibrations incluses dans cette zone sont jugées non satisfaisantes pour des opérations constantes et prolongées. Généralement, la machine peut être utilisée

pour une période de temps limitée sous ces conditions jusqu'à ce que des mesures de correction puissent être adoptées.

- Zone D : Les valeurs de vibration incluses dans cette zone sont normalement jugées suffisamment graves pour causer des dégâts à la machine.

Ces évaluations permettent à l'opérateur de diagnostiquer les conditions vibratoires d'une machine et offrent des instructions claires concernant toute mesure nécessaire.

Limites de la zone d'évaluation pour la vitesse de vibration V_{eff} pour les mesurages sur des parties non tournantes par la norme ISO 10816-4.

- Limite de la zone A/B $V_{\text{eff}} = 4,5 \text{ mm/s}$
- Limite de la zone B/C $V_{\text{eff}} = 9,3 \text{ mm/s}$ (alarme)
- Limite de la zone C/D $V_{\text{eff}} = 14,7 \text{ mm/s}$ (procédure de panne)

Les valeurs s'appliquent aux vibrations du boîtier dans la direction verticale sous des conditions de fonctionnement en régime permanent et à la vitesse de fonctionnement.

Tout changement survenu dans l'amplitude de vibration par rapport à une valeur de référence préétablie constitue un second critère d'évaluation. Si une intensification rapide ou instantanée des vibrations était observée, mener des enquêtes pertinentes conjointement avec le fabricant, même si le niveau général demeure en dessous des valeurs d'alarme $V_{\text{eff}} = 9,3 \text{ mm/s}$. Si un changement dans l'amplitude de vibration était accompagné de bruits retentissants provenant de la TG, éteindre le turbogénérateur immédiatement même si l'exactitude de l'affichage est en doute. De tels bruits sont révélateurs de dégâts mécaniques (dégât des aubes ou dégâts causés par un frottement excessif), [34].

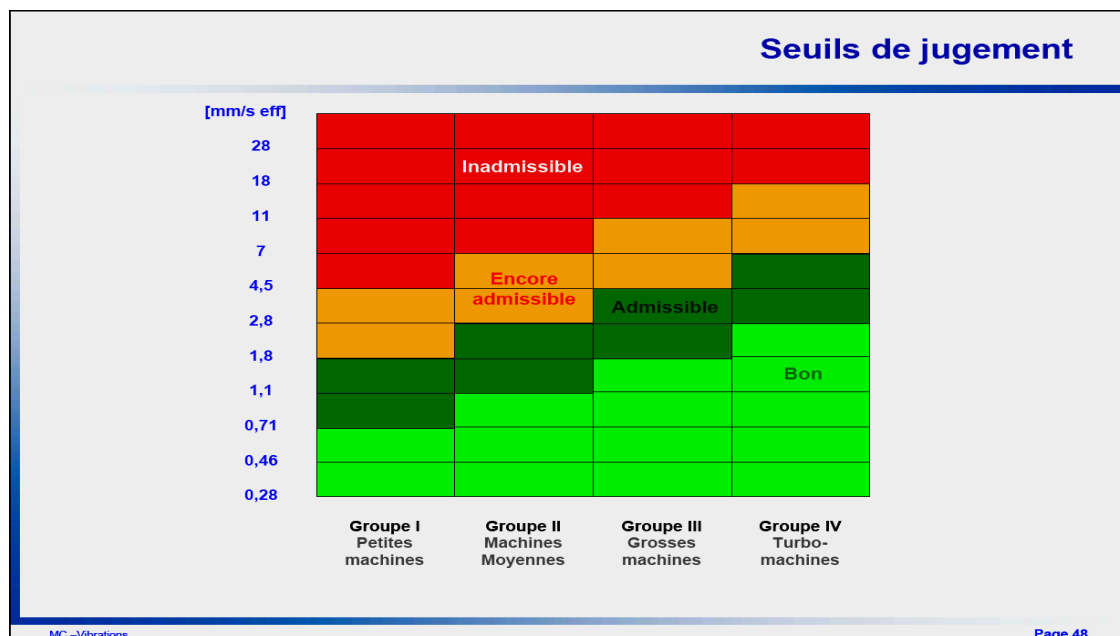


Figure III.2. Seuil de Jugement de vibration,[34].

III.4. Environnement de Travail

La turbine à gaz de type V94.3A de fabrication ANSALDO est de type industriel à un seul arbre (un seul rotor) et un seul corps avec une puissance minimale 215 MW. Est capable de faire fonctionner les générateurs dans les centrales. On peut utiliser ce type des turbines dans des cycles combiné ou simple. Elles peuvent fonctionner avec des combustibles liquides, tels que les huiles combustibles légères, ou les combustibles gazeux ayant des valeurs spécifiques différentes, comme le gaz naturel.

Le rotor du turbo-alternateur des groupes de la centrale M'sila III (V94.3A) se repose sur quatre (04) paliers lisses (palier à huile). Les vibrations sont mesurées près du palier de la turbine et près du palier du compresseur à l'aide des capteurs de vibrations 11MBD11CY101 et 11MBD11CY102 (palier de turbine), 11MBD12CY101 et 11MBD12CY102 (palier de compresseur), et 11MKD10/20CY101 et 11MKD10/20CY102 (palier de générateur). Un module processeur calcule la vitesse de vibration effective de chaque signal. Deux canaux équivalents sont utilisés pour mesurer la vibration du carter : ils sont désignés comme « Canal 1 » et « Canal 2 », pour des liens logiques additionnels. Si au moins un des deux canaux indique que la limite est dépassée, l'alarme « Alarme de vibration du palier » est déclenchée. Si un canal indique que la limite supérieure est dépassée, la procédure de panne commence et la turbine à gaz est éteinte.

III.5. Collection et Prétraitement du DATA:

Afin d'évaluer un modèle d'apprentissage automatique pour la détection et classification des défauts de la turbine à gaz, nous utilisons la collection du DATA acquise à partir de l'analyseur de vibrations OROS 35. Archivée en date du 25/08/2020 et 03/06/2022 des groupes turbines à gaz de la centrale M'sila III.

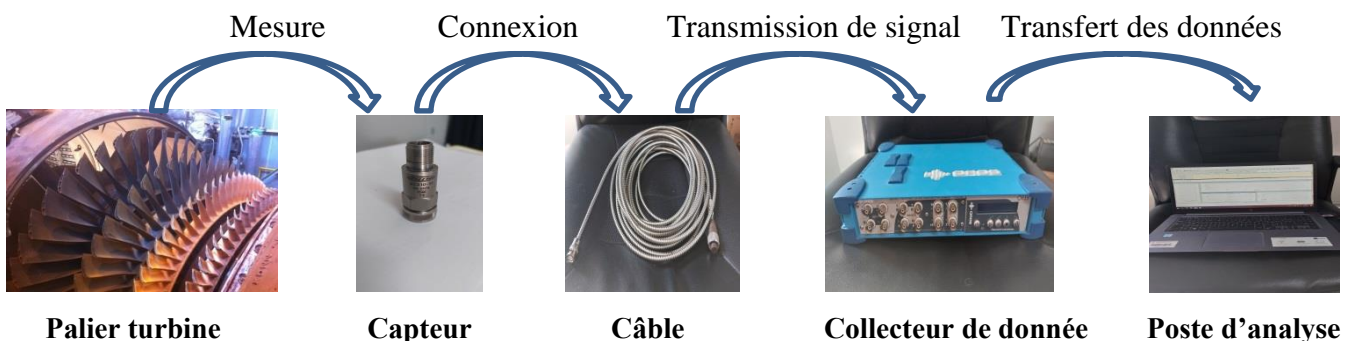


Figure III.3. Chaîne de mesure et collecte des données

A partir du de programme ci-dessous sur Python on peut tracer les données de vibration collectées.

```
# Test for the first file
dataset = pd.read_excel('/content/data22.xlsx')
ax = dataset.plot(figsize = (24,6), title= " signal vibratoire brute" , legend = True)
ax.set(xlabel="enchantions", ylabel="vibration en mm/s")
plt.show()
```

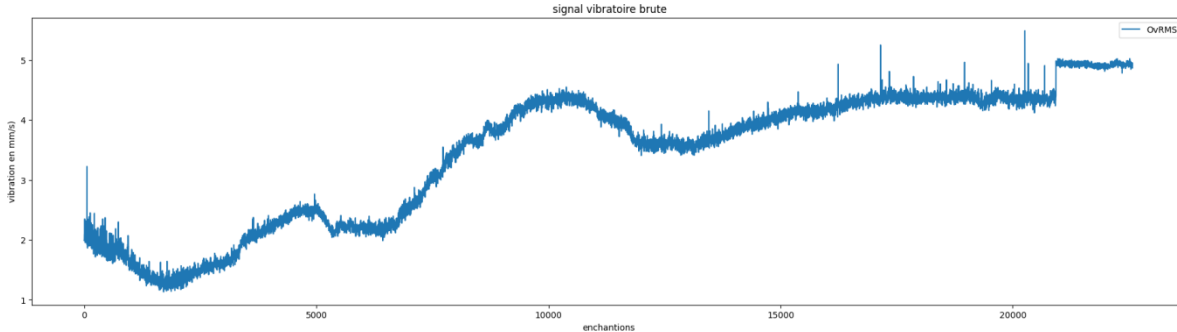


Figure III.4. Signal de vibration des données collectées

Cette collection contient énormément de données manquantes (case vide, valeur hors échelle ...). Par conséquent nous avons les nettoyés avec les différentes techniques que nous avons décrites en ANNEX I, beaucoup de paramètres donc sont supprimés.

Puis, nous avons divisé le DATA (22000 valeurs) en sous-groupe de données (100 valeurs par sous-groupe) on obtient 220 enchantions. Et ce pour extraction des indicateurs.

III.6. Extraction des indicateurs

III.5.1. Kurtosis

Longueur des queues d'une distribution de signal, ou de manière équivalente, degré de tendance du signal aux valeurs aberrantes. Le développement de défauts peut augmenter le nombre de valeurs aberrantes, et donc augmenter la valeur de la métrique d'aplatissement.

$$K = \frac{1}{N} \sqrt{\sum_{i=1}^N \frac{(x_i - \bar{x})^4}{\alpha^4}} \dots\dots\dots \text{III.1}$$

III.5.2. Skewness

Asymétrie d'une distribution de signal. Les défauts peuvent avoir un impact sur la symétrie de la distribution et donc augmenter le niveau d'asymétrie.

$$SK = \frac{1}{N} \sqrt{\sum_{i=1}^N \frac{(x_i - \bar{x})^3}{\alpha^3}} \dots\dots\dots \text{III.2}$$

III.5.3. Facteur de forme

RMS divisé par la moyenne de la valeur absolue. Le facteur de forme dépend de la forme du signal tout en étant indépendant des dimensions du signal.

$$F_c = \frac{x_{rms}}{\bar{x}} \dots\dots\dots \mathbf{III.3}$$

III.5.4. Impulsion

Comparez la hauteur d'un pic au niveau moyen du signal.

$$I_m = \frac{\text{max value}}{\bar{x}} \dots\dots\dots \mathbf{III.4}$$

III.5.5. RMS

L'extraction de caractéristiques fait référence au processus de sélection ou d'extraction d'informations ou de caractéristiques pertinentes à partir de données brutes. Dans le cadre de l'extraction de caractéristiques utilisant RMS (Root Mean Square), RMS peut être utilisé comme caractéristique pour représenter certaines caractéristiques d'un signal ou d'une série temporelle.

$$x_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x)^2} \dots\dots\dots \mathbf{III.5}$$

III.5.6. Moyenne absolue

La "moyenne absolue" (Absolute Mean en anglais) peut être utilisée comme une caractéristique dans l'extraction de caractéristiques d'un ensemble de données. Elle représente la moyenne des valeurs absolues des points de données.

$$X = \frac{1}{N} \sum_{i=1}^N |x_i| \dots\dots\dots \mathbf{III.6}$$

III.5.7. Écart type

L'écart type (standard deviation en anglais) peut être utilisé comme une caractéristique dans l'extraction de caractéristiques d'un ensemble de données. L'écart type mesure la dispersion des valeurs par rapport à la moyenne et peut donner une indication de la variabilité des données.

$$\alpha = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - x)^2} \dots\dots\dots \mathbf{III.7}$$

III.5.8. Facteur de dégagement :

Valeur de crête divisée par la valeur moyenne au carré des racines carrées des amplitudes absolues. Pour les machines tournantes, cette caractéristique est maximale pour les roulements sains et continue à diminuer pour la bille défectueuse, la bague extérieure défectueuse et la bague intérieure défectueuse respectivement. Le facteur de jeu a la capacité de séparation la plus élevée pour les défauts de bague intérieure défectueuse

$$x_{clear} = \frac{x_p}{(\frac{1}{N} \sum_{i=1}^N \sqrt{|x_i|})^2} \dots\dots\dots \text{III.8}$$

III.5.9. Entropie

L'entropie est une mesure de l'incertitude ou de la complexité d'un ensemble de données. Elle peut être utilisée comme une caractéristique dans l'extraction de caractéristiques pour représenter la distribution ou la variabilité des valeurs dans un ensemble de données.

$$H(X) = - \sum_{i=1}^N P(c) \log P(x_i) \dots\dots\dots \text{III.9}$$

III.5.10. Peak to Peak

Le "Peak to Peak" (Pic à Pic en français) peut être utilisé comme une caractéristique dans l'extraction de caractéristiques d'un ensemble de données. Il représente la différence entre la valeur maximale et la valeur minimale d'un signal ou d'une série de données.

$$X_p = \text{Maxvalue} - \text{Minvalue} \dots\dots\dots \text{III.10}$$

III.5.11. Facteur de crête

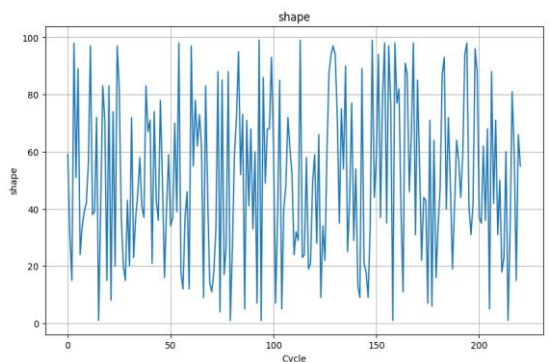
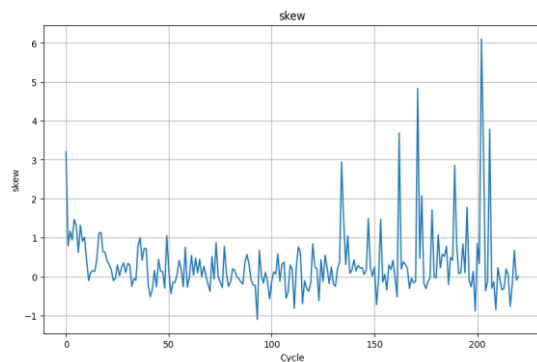
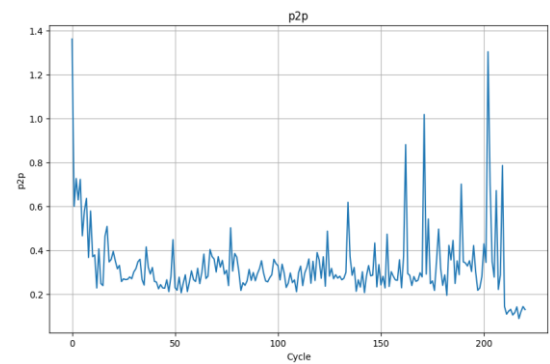
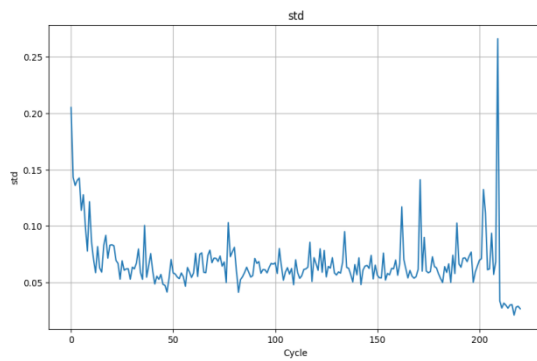
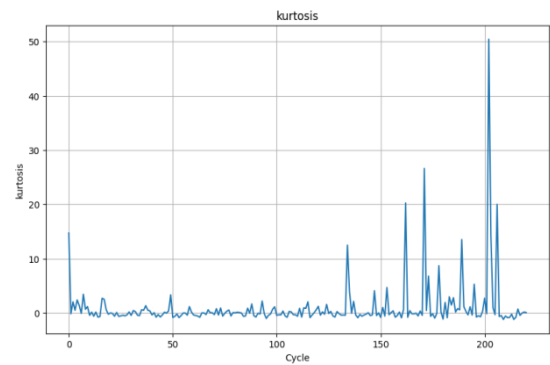
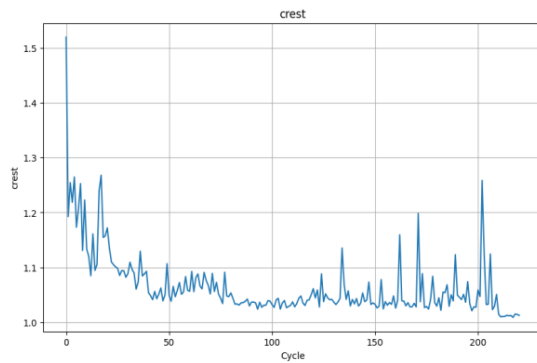
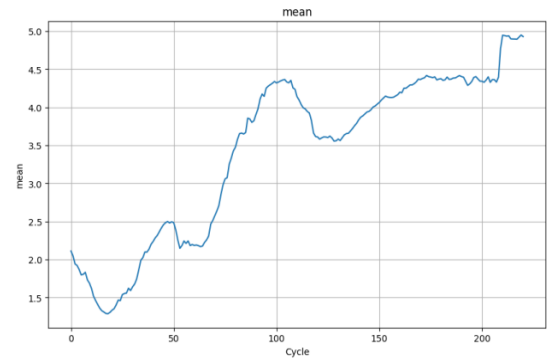
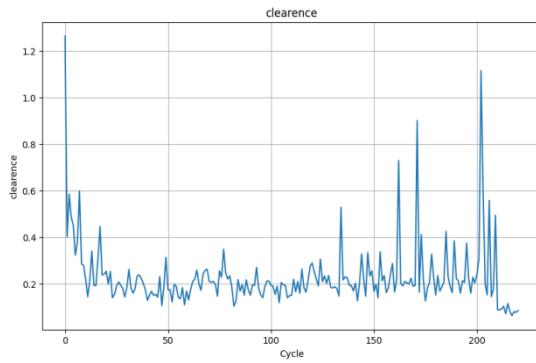
valeur de crête divisée par la valeur efficace. Les défauts se manifestent souvent d'abord par des changements dans la crête d'un signal avant de se manifester dans l'énergie représentée par la moyenne quadratique du signal. Le facteur de crête peut fournir une alerte précoce pour les défauts lorsqu'ils se développent pour la première fois.

$$x_{clear} = \frac{\text{maxvalue}}{x_{rms}} \dots\dots\dots \text{III.11}$$

	mean	std	skew	kurtosis	entropy	rms	max
0	2.111428	0.204943	3.198635	14.725673	4.600846	2.121351	3.224367
1	2.049577	0.143119	0.785921	-0.138476	4.602771	2.054568	2.450076
2	1.941812	0.135903	1.171087	2.052676	4.602779	1.946562	2.442009
3	1.921857	0.140407	0.941918	0.501660	4.602555	1.926979	2.348376
4	1.868157	0.142498	1.472440	2.403298	4.602356	1.873584	2.369875

	P2p	crest	clearence	shape	impulse
0	1.360490	1.519959	1.264582	59	3.224367
1	0.600095	1.192502	0.403449	31	2.450076
2	0.726395	1.254524	0.585564	15	2.442009
3	0.629528	1.218683	0.487286	98	2.348376
4	0.722625	1.264889	0.452436	51	2.369875

Résultat d'extraction des indicateurs



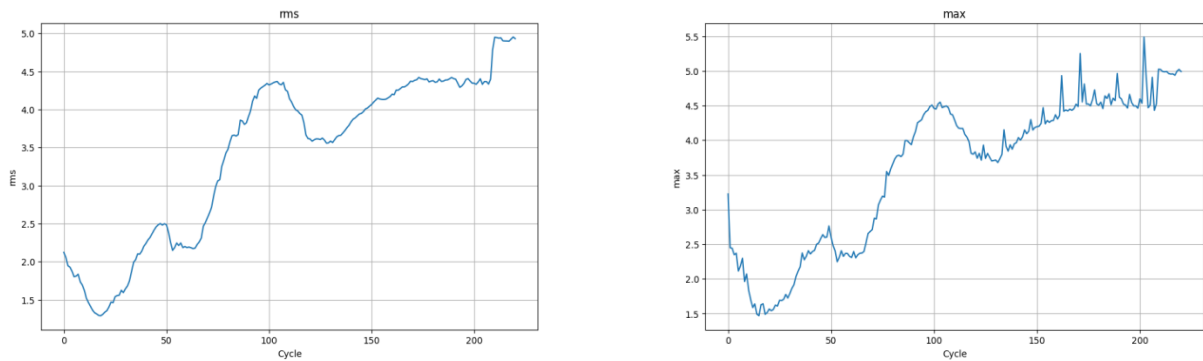


Figure III.5. Représentation graphique des indicateurs

III.6. Phase d'apprentissage

Après le prétraitement, nous avons divisé les données en deux (02) parties (70 % pour la l'apprentissage, 30 % pour la validation (test) et cela en utilisant la fonction split de la bibliothèque Scikit-Learn `''from sklearn.model_selection import train_test_split X_train, X_test,y_train,y_test=train_test_plit(X, y, test_size=0.33, random_state=42)''`.

Puis les données sont passées aux modèles pour apprentissage. Où nous avons utilisé pas mal d'algorithmes de classification. Et pour les évaluer on utilise quelque critère de validation pour sélectionner le meilleur modèle de classification.

III.7 Résultats de simulation

III.7.1 Résultats d'apprentissage des Modelés :

Les modèles utilisés, leurs programmes et leurs résultats sont comme suit :

1. Le Réseau de Neurone Multicouches (MLP) :

La structure de réseau MLP utilisé est comme suite :

Une couche caché qui contient 100 neurones, avec l'utilisation de la fonction sigmoïde comme une fonction d'activation, l'apprentissage ce faite pour 1000 itérations avec un pas d'apprentissage égale à 0.001. Le syntaxe de cet algorithme en python avec l'utilisation de la bibliothèque SKTLEARN est le suivant :

```
modelmlp=MLPClassifier(hidden_layer_sizes=(100), activation=' sigmoïde
', *, solver='adam', alpha=0.0001, batch_size='auto', learning_rate='constant'
, learning_rate_init=0.001)
modelmlp.fit(X_train, y_train)
preds = modelmlp.predict(X_test)
```

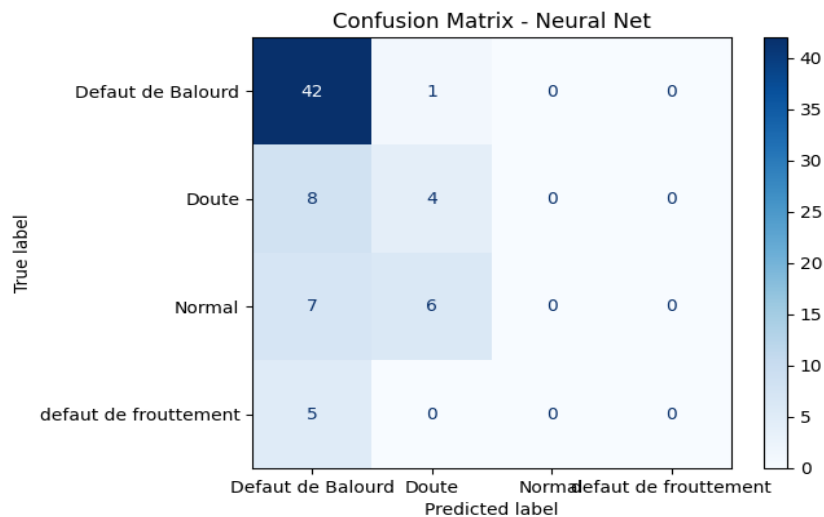


Figure III.6 Matrice de confusion des défauts pour le classificateur de ‘MLP’.

Nous avons utilisé une matrice de confusion pour montrer visuellement la classification. des résultats de prédiction d'un problème de classification sont compilés dans une matrice de confusion. Les valeurs de comptage sont utilisées pour résumer le nombre de Prédications inexactes pour chaque classe. La figure III.9 montre la matrice de confusion obtenue après ajustement du modèle MLP proposé. L'axe des x représente les défauts réels et l'axe des y représente les défauts annoncés. La diagonale de la matrice montre les vraies prédictions.

2. Modèle ‘Arbre de décision’ :

```

modeld=DecisionTreeClassifier(max_depth=5)
11modeld.fit(X_train, y_train)
preds= modeld.predict(X_test)

```

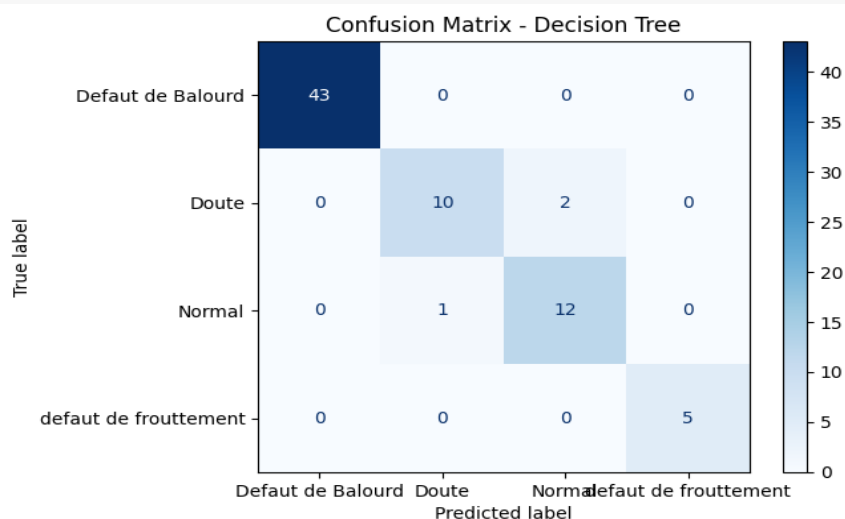


Figure III.7 Matrice de confusion des défauts pour le classificateur de ‘Arbre de décision’

3. Modèle ‘Random Forest’

```

modeld=RandomForestClassifier(max_depth=5, n_estimators=10, max_features=1),

modeld.fit (X_train, y_train)
preds= modeld.predict(X_test)
    
```

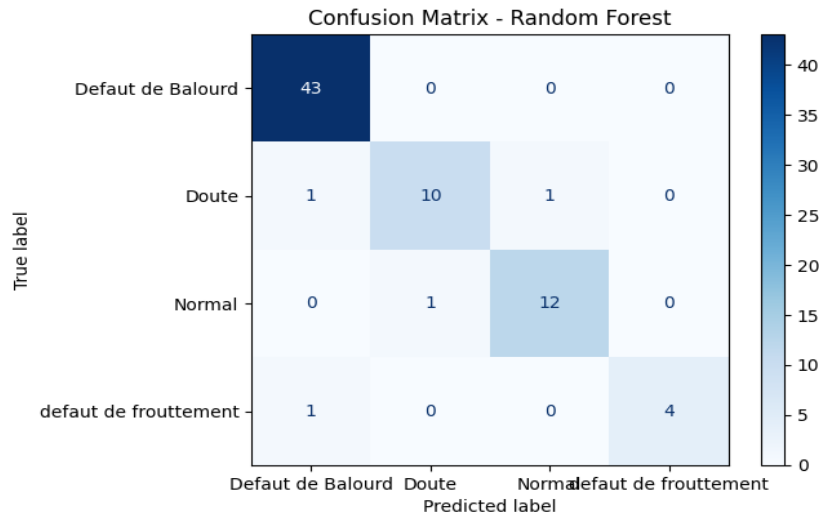


Figure III.8 Matrice de confusion des défauts pour le classificateur de ‘Random Forest’

4. Modèle ‘Linear SVM’

```

model= Linear.SVMClassifier()
model.fit(X_train, y_train)
preds = model.predict(X_test)
    
```

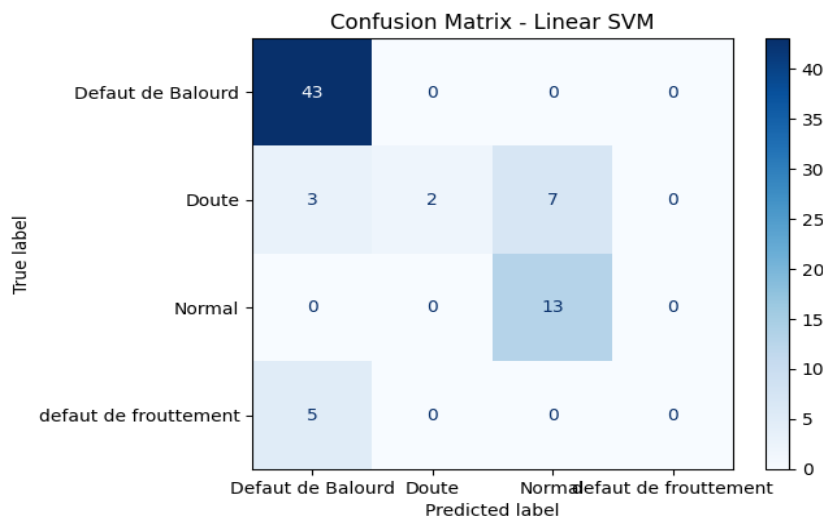


Figure III.9 Matrice de confusion des défauts pour le classificateur de ‘Linear SVM’

5. Modèle “KNeighbors”

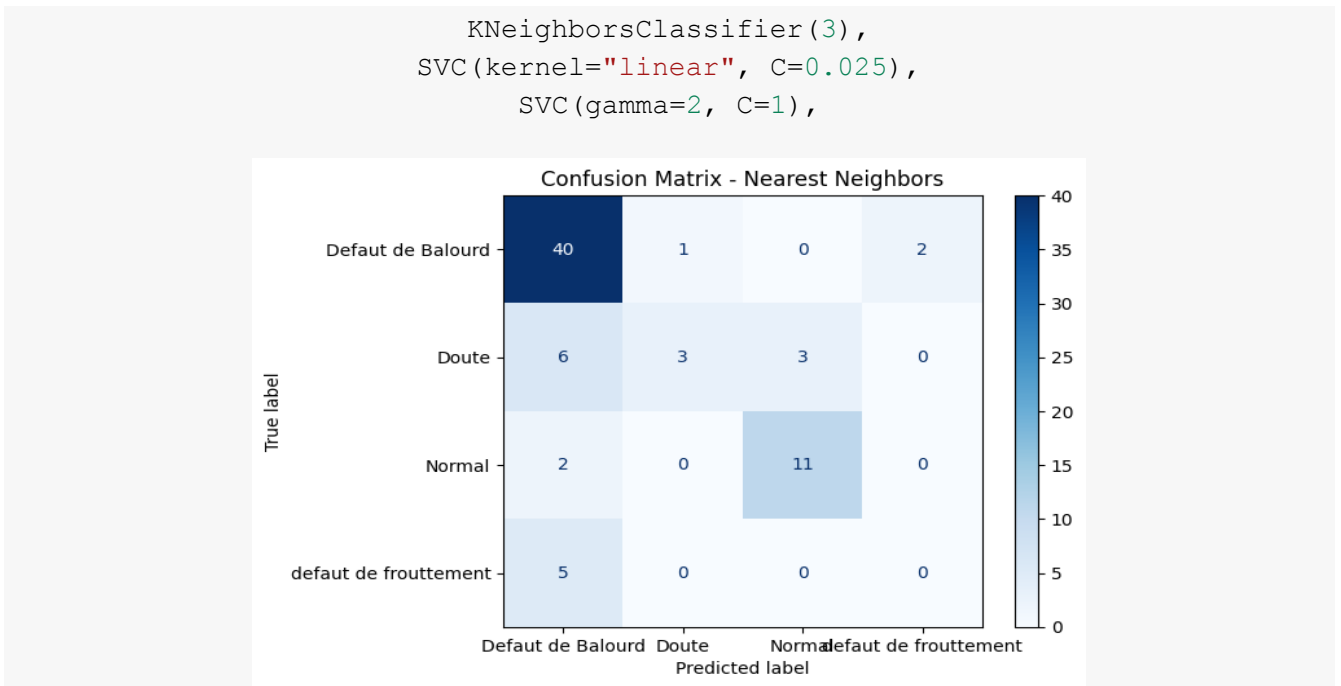


Figure III.10 Matrice de confusion des défauts pour le classificateur de ‘KNeighbors’

6. Modèle “Naive Bayes”

```

Model6= GaussianNB(),
Model6.fit (X_train, y_train)
preds= modeld.predict(X_test)
    
```

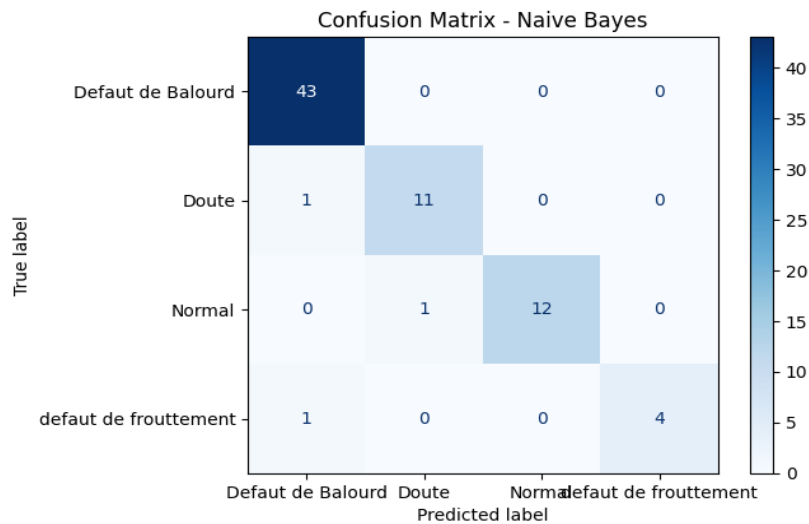


Figure III.11 Matrice de confusion des défauts pour le classificateur de ‘Naive Bayes’

7. Modèle ‘XGboost’

```
xgb_cl = xgb.XGBClassifier(use_label_encoder=False)
xgb_cl.fit(X_train, y_train)
preds = xgb_cl.predict(X_test)
```

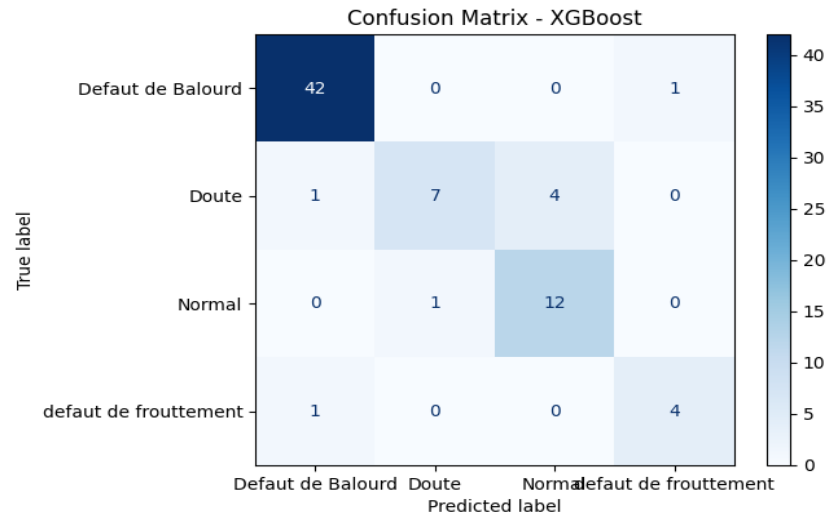


Figure III.12 Matrice de confusion des défauts pour le classificateur de ‘XgBoost’

8. Modèle ‘Adaboost’

```
model= AdaBoostClassifier()
model.fit(X_train, y_train)
preds = model.predict(X_test)
```

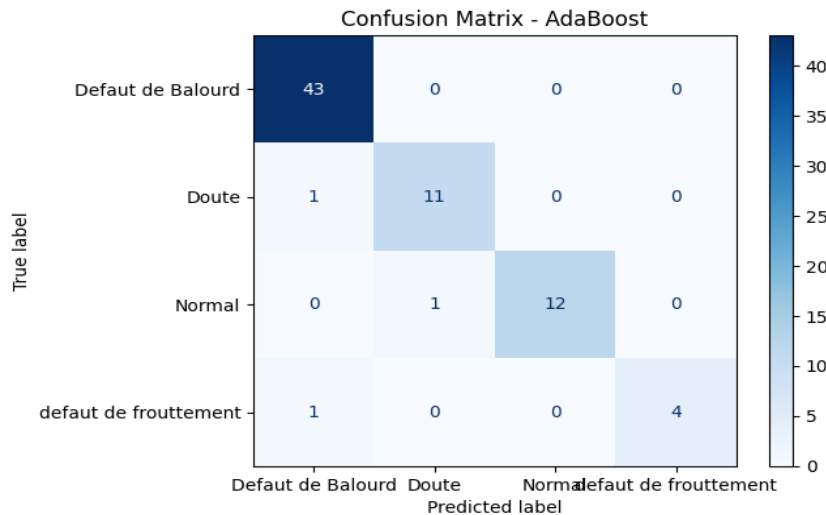
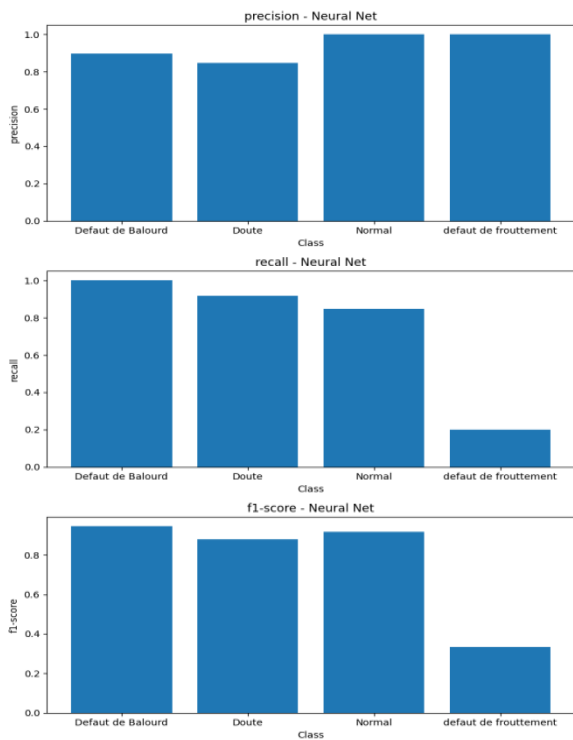


Figure III.13 Matrice de confusion des défauts pour le classificateur de ‘AdaBoost’

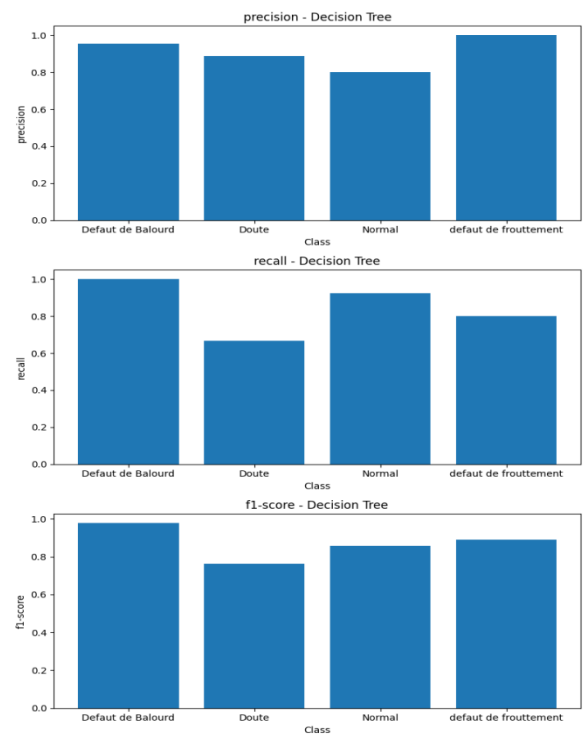
III.7.2 Evaluation des Modèles

L’évaluation des modèles proposés à l’apprentissage se fait selon les critères que nous avons définis au CHAPITRE II de ce mémoire. Il s’agit de la précision, recall et score. Et on va les représenter sous forme de Bargraphe pour chaque modèle comme suit :

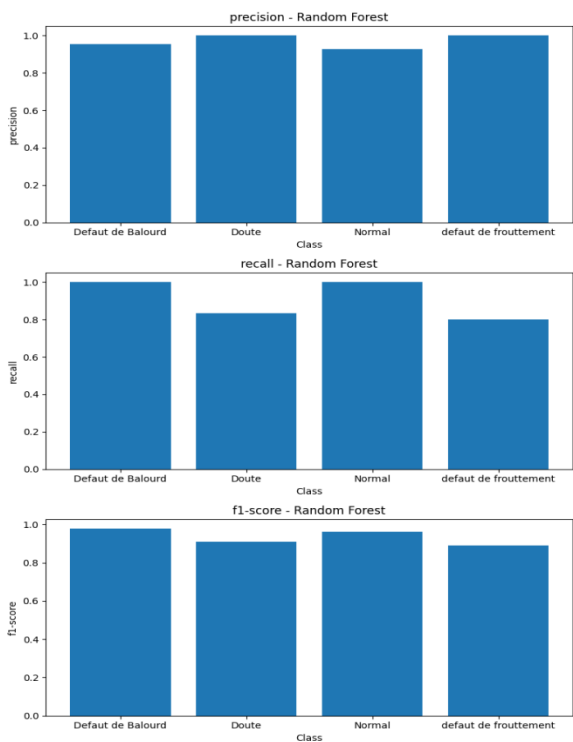
1. Neurone Multicouches (MLP)



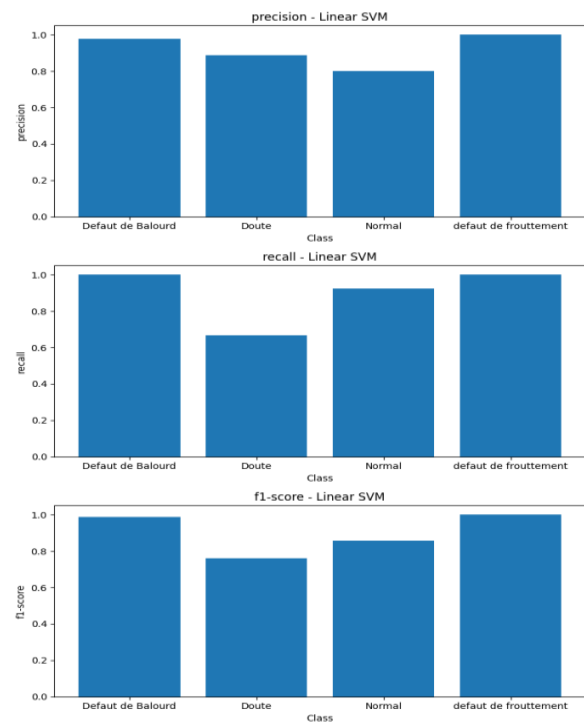
2. Arbre de décision



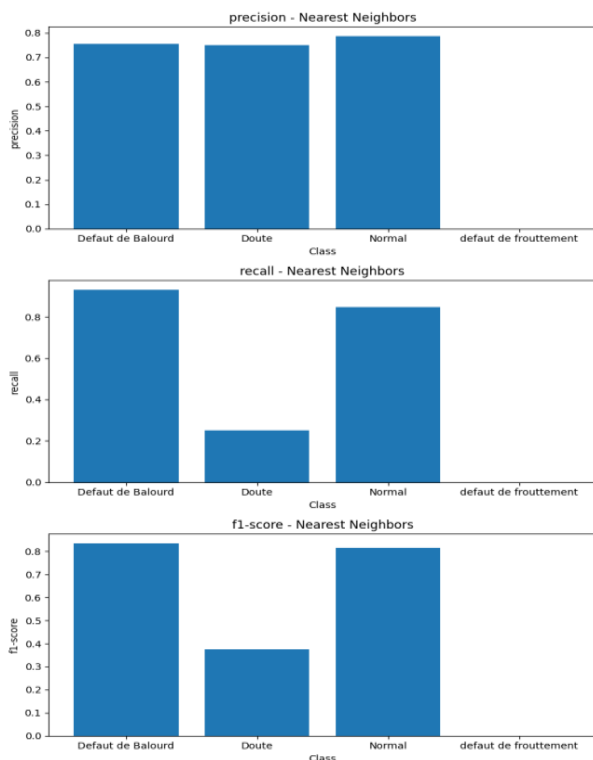
3. Random Forest



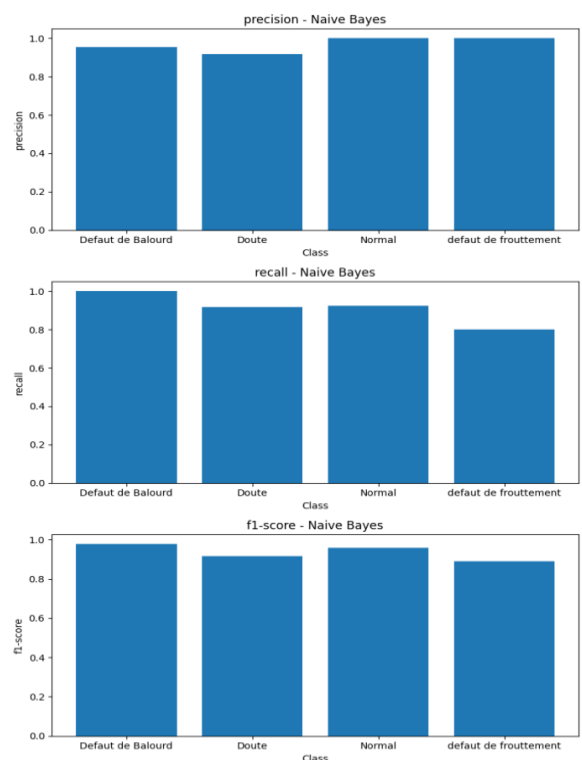
4. Linear SVM



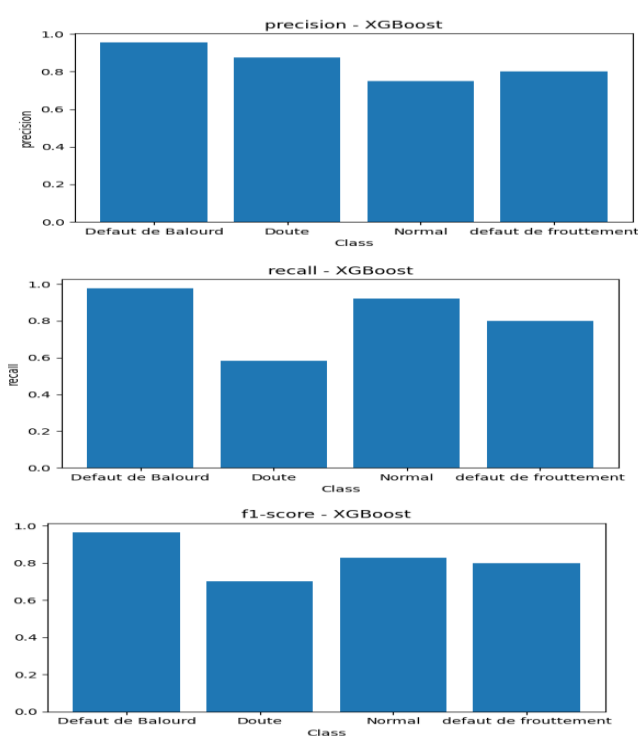
5. KNeighbors



6. Naive Bayes



7. XgBoost



8. AdaBoost

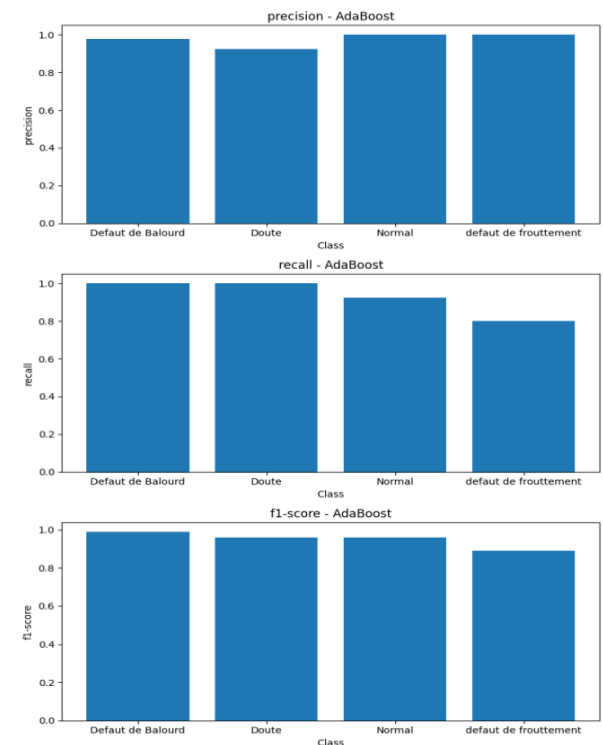


Figure III.14 Résultats d'évaluation des Modèles proposés

III.7.3 Etude Comparative et Choix d'Algorithme :

Les résultats sur la classification des défauts des turbines à gaz sont données dans les Figures(III.9 –III.16). D'après le TABLEAU III.1 on peut voir que pour notre ensemble de données, les résultats les plus précis sont reçus par le classificateur de **Adaboost**. Cette méthode donne 97,26 % du score, Cela signifie que le modèle formé par **Random Forest** et **Naive Bayes** donnera des bonnes score : 95,89%.

Le Tableau ci-dessous montre le résultat final obtenu :

Algorithme	Arbre de décision	XgBoost	AdaBoost	Neural Net	Linear SVM	Random forest	Naive Bayes	KNeighbors
Accuracy	91.78%	89.04%	97.26%	93.15%	93.15	95.89%	95.89%	73.97%

Tableau III.1 : Comparaison des résultats des algorithmes

Le tableau III.1 montre un récapitulatif des résultats des huit modèles entraînés sur la même collection de données. Et pour commencer, nous observons très nettement que l'on a obtenu de meilleurs résultats avec le modèle "AdaBoost".

Concernant les modèles "Random forest" et "Naive Bayes" nous obtenons des résultats assez similaires pour les trois mesures de performances.

Donc dans ce travail comme résultat final on adopte le modèle "AdaBoost" comme modèle type pour la détection et classification des défauts des vibrations de la turbine à gaz.

III.8.Conclusion

Dans ce dernier chapitre, nous avons présenté dans un premier temps, la méthode d'analyse et les critères d'évaluation des vibrations utilisés actuellement dans notre ouvrage de travail de ce mémoire (Turbine à Gaz V94.3A de la centrale de M'sila) puis les différentes bibliothèques python utilisées dans le domaine de l'apprentissage automatique.

Par la suite, nous avons montré la manière d'effectuer les différentes étapes de prétraitement des données de vibrations, suivie par la création des données d'entraînement et de test. Puis, nous avons vu l'apprentissage et le test des modèles proposés. Enfin, nous avons effectué une comparaison des résultats de ces trois modèles.

La comparaison des résultats des huit modèles nous a permis de déduire le modèle adéquat à la détection et classification des défauts de vibrations de la TAG où nous a donné un résultat très satisfaisant.

CONCLUSION GENERALE

Dans cette étude, nous avons exploré la détection et la classification des défauts des turbines à gaz en utilisant des techniques d'apprentissage automatique. Les turbines à gaz jouent un rôle crucial dans diverses industries, et l'identification en temps opportun des défauts est essentielle pour garantir leur fonctionnement fiable et efficace. Nous avons démontré l'importance et la puissance de l'intelligence artificielle et de l'apprentissage automatique dans la réalisation de cet objectif.

En exploitant les capacités de l'IA et de l'apprentissage automatique, nous avons développé des systèmes robustes de détection des défauts capables d'identifier et de classer avec précision les défauts courants des turbines à gaz. Grâce à l'utilisation d'algorithmes avancés tels que les arbres de décision, les machines à vecteurs de support, les réseaux neuronaux et les méthodes d'ensemble, nous avons obtenu des niveaux élevés de précision et de fiabilité dans la détection des défauts.

Dans le chapitre 1, nous avons fourni une description détaillée des turbines à gaz, de leur fonctionnement et des défauts potentiels pouvant survenir. Comprendre la nature et l'impact de ces défauts est essentiel pour le développement de systèmes efficaces de détection et de classification.

Le chapitre 2 a introduit les fondamentaux des algorithmes d'apprentissage automatique et a mis en évidence leur importance dans la détection des défauts. Nous avons exploré différentes techniques et algorithmes, en discutant de leurs forces et de leurs limitations. Ce chapitre a posé les bases de l'application de l'apprentissage automatique dans la détection des défauts des turbines à gaz.

Dans le chapitre 3, nous avons présenté la méthodologie, la configuration expérimentale et les résultats de l'application des algorithmes d'apprentissage automatique pour détecter et classer les défauts dans les turbines à gaz. Nous avons collecté des données du monde réel, les avons prétraitées, avons conçu des caractéristiques pertinentes et avons implémenté les algorithmes sélectionnés. Les résultats ont démontré l'efficacité de l'apprentissage automatique dans la détection des défauts, avec une grande précision et un nombre minimal de faux positifs.

Cette recherche montre que l'IA et les techniques d'apprentissage automatique peuvent considérablement améliorer les capacités de détection des défauts des turbines à gaz. L'identification précoce des défauts permet une maintenance proactive, réduisant les temps d'arrêt et améliorant l'efficacité opérationnelle globale.

En conclusion, cette étude met en évidence l'importance et la puissance de l'intelligence artificielle et de l'apprentissage automatique dans la détection et la classification des défauts des turbines à gaz. L'application de ces techniques a démontré leur efficacité pour identifier et classifier avec précision les défauts, permettant ainsi des interventions de maintenance en temps opportun et évitant les défaillances potentielles. Les résultats de cette recherche contribuent à l'avancement des méthodologies de détection des défauts dans les turbines à gaz et fournissent des informations précieuses pour les praticiens de l'industrie.

Des recherches supplémentaires peuvent être menées pour explorer des algorithmes d'apprentissage automatique plus avancés, intégrer des sources de données supplémentaires et améliorer la précision de la classification des défauts. De plus, le développement de systèmes de surveillance en temps réel intégrant l'IA et l'apprentissage automatique peut améliorer davantage la maintenance proactive des turbines à gaz. En adoptant ces avancées, les industries peuvent garantir un fonctionnement continu et optimal des turbines à gaz, ce qui se traduit par une fiabilité accrue, des économies de coûts et une amélioration de la sécurité.

Références Bibliographiques

- [1] **N.E Boumediene et A. Nourine**, « Calcul du rendement de turbine à gaz par la méthode de Chambadal à différentes températures », Mémoire de Master, Université Abdelhamid Ibn Badis Mostaganem, 25 juin 2020.
- [2] **L. BOUSSOUAR et S. NEDJIMI**, « Etude de la transformation du groupe n°1 de la centrale de Cao Djenet en cycle combiné », Mémoire de Master, Université Mouloud Mammeri de Tizi-Ouzo, 2015/2016.
- [3] **Ansaldo Energia**, « Volume de Formation », Projet de la centrale de MS III, 2007.
- [4] **G. J.**, « Survey of model based-based failure detection and isolation in complex plants », IEEE Control System Magazine, pp. 3-11, 1988.
- [5] **S.G. Gertler.j.j**, « A new structural framework for parity equation-based failure detection and isolation », Automatica Journal, vol. 26, n° 12, pp. 381-388, 1990.
- [6] **Landolsi Foued**, « cours de techniques de surveillance ».
- [7] **A. H. a. K. A. Benrabe Djaidir**, « Performance control of gas turbines », Turbomachinery International Journal, pp. 10-14, 2015.
- [8] **Djadair Benrabe**, « Modélisation et simulation des vibrations du rotor BP d'une turbine à gaz MS 5002B, et l'effet de réparation de masse sur le roulement à bille », The first international conference On applied automation and industrial diagnostics (ICAAID 2015), Djelfa on 29-30 March 2015.
- [9] **ABBAS Faouzi et HARRAT Karim**, « Diagnostic vibratoire des défauts de types chocs par les méthodes temps -échelle », Université Mouloud Mammeri de Tizi Ouzou, 2019/2020.
- [10] **ELFAR Youssouf**, « Diagnostic Des Défauts D'une Machine Tournante Par L'analyse Vibratoire », Mémoire Master II, l'UNIVERSITE de M'sila, 2011/2012.
- [11] **Ansaldo Energia**, « Manuel de la Turbine », Projet sonelgaz M'sila, Contrat 2007.
- [12] **LANDOLSIA**, « Cours De Technique De Surveillance », 2010.
- [13] **DEBBAH YOUNES**, « Mécanique De Vibration ».
- [14] **M. Zemouri**, « contribution à la surveillance des systèmes de production à l'aide des réseaux de couronnes dynamique application à la e-maintenance », Thèse de Doctorat, France, 2003.
- [15] **Ahmed KEROUCHA et Youcef BENHAMIDOUCHE**, « Prediction Of Cement Fineness Using Machine Learning Approaches », Mémoire Master, Université de M'sila 2020/2021.
- [16] **R. J. Solomonoff**, « A new method for discovering the grammars of phrase structure languages. In COMMUNICATIONS OF THE ACM (Vol. 2, No. 7, pp. 20-20), vol. Vol 2 », 1515. BROADWAY,

NEW YOURK, 1959.

- [17] **Z. Pawlak**, **Rough sets**, «Theoretical aspects of reasoning about data, vol. Vol. 9», USA: Springer Science & Business Media, 2012.
- [18] **F. Rosenblatt**, «The perceptron, a perceiving and recognizing automaton Project Para», Cornell Aeronautical Laboratory., Buffalo, Neu York, 1957.
- [19] **J. L. R. D. E. & P. R. G. McClelland**, «Parallel distributed processing (Vol. 2, pp. 20-21).», Cambridge, MA, USA: MIT press, 1986.
- [20] <https://datascientest.com/machine-learning-tout-savoir>.
- [21] **A. Géron**, «Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems», USA: O'Reilly Media, 2019.
- [22] **C. Nikki**, « Supervised vs Unsupervised Machine Learning», 2017.
- [23] <https://www.talend.com/fr/resources/etapes-machine-learning/>
- [24] <https://brightcape.co/5-apprentissage-supervise/>
- [25] <https://www.dataanalyticspost.com/Lexique/svm>
- [26] **Steeven Janny.Solal Nathan, winqi SHU-Quartier**, «Introduction A L'apprentissage Automatique» Edité le 24/05/2022.
- [27] **FATIMA AIT MAHAMMED**, «Approches D'apprentissage Automatique Pour La Détection Du Spam Web : Exploration De Diverses Caractéristiques» Mémoire fin d'Etude, UNIVERSITÉ DU QUÉBEC À MONTRÉAL, Mars 2018.
- [28] **Hacene BELLAHMER**, «Implémentation et évaluation d'un modèle d'apprentissage automatique pour l'estimation de la valeur marchande de propriétés immobilières», Mémoire de fin d'études 2019/2020.
- [29] <https://www.xlstat.com/fr/solutions/fonctionnalites/classifieur-bayesien-naif>.
- [30] <https://fr.linedata.com/quest-ce-que-lapprentissage-non-supervise>
- [31] <https://fr.linedata.com/principaux-algorithmes-dapprentissage-non-supervise>
- [32] <https://inside-machinelearning.com/en/recall-precision-f1-score-simple-metric-explanation-machine-learning/>
- [33] **Touam Fouad & Hadj Mihoub Sidi Moussa Fayçal**, «Détection et localisation des défauts des capteurs», Mémoire de Projet de Fin d'Étude, Université SAAD DAHLAB de BLIDA, 2012/2013.
- [34] **Adsaldo Energia**, V94.3a Gas Turbine Operating And Maintenance Manual (O&M) M'sila.
- [35] **Patrick Fuchs et Pierre Poulain**, «Cours de Python», version du 29 août 2022, Université Paris Cité, France.

- [36] **Balta-Ozkan, N., Davidson, R., Bicket, M., & Whitmarsh, L.** (2013). Social barriers to the adoption of smart homes. *Energy Policy*, *63*, 363–374.
- [37] **Mohamed Abd Elmoumen DJABALLAH**, «Système de prédiction de la consommation d'énergie basé Deep Learning», Mémoire de Fin d'études Master, Université de 8 Mai 1945 – Guelma – sep 2021.

Annexe

Le présent Annexe présente **Environnement Logiciel** et les différentes bibliothèques du colab et le langage de Programmation Python utilisées dans ce mémoire :

Le logiciel PYTHON

Le langage de programmation Python a été créé en 1989 par Guido van Rossum, aux Pays-Bas. Le nom Python vient d'un hommage à la série télévisée Monty Python's Flying Circus dont G. van Rossum est fan. La première version publique de ce langage a été publiée en 1991.

Ce langage de programmation présente de nombreuses caractéristiques intéressantes : - Il est multiplateforme. C'est-à-dire qu'il fonctionne sur de nombreux systèmes d'exploitation : Windows, Mac OS, Linux, Android, iOS, depuis les mini-ordinateurs Raspberry Pi jusqu'aux supercalculateurs. - Il est gratuit. Vous pouvez l'installer sur autant d'ordinateur que vous voulez. - C'est un langage de haut niveau. Il demande relativement peu de connaissance sur le fonctionnement d'un ordinateur pour être utilisé. - C'est un langage interprété. Le programme n'a pas besoin de compiler son programme pour pouvoir l'utiliser, contrairement à des langages comme le C ou le C++. - Il est orienté objet. C'est-à-dire qu'il est possible de créer en Python des entités qui ont un sens dans le monde réel (une cellule, une protéine, un atome) avec un certain nombre de fonctionnements et d'interactions. - Il est relativement simple à prendre en main. - Enfin, il est très utilisé en bio-informatique et plus généralement en analyse de données.

Toutes ces caractéristiques font que Python est désormais enseigné dans de nombreuses formations, depuis l'enseignement secondaire jusqu'à l'enseignement supérieur.

Python est devenu ces dernières années le langage de programmation le plus employé par les informaticiens. Ce langage s'est propulsé en tête de la gestion d'infrastructure, d'analyse de données ou dans le domaine du développement de logiciels. Python permet notamment aux développeurs de se concentrer sur ce qu'ils font plutôt que sur la manière dont ils le font. Il a libéré les développeurs des contraintes de formes qui occupaient leur temps avec les langages plus anciens. Ainsi, développer du code avec Python est plus rapide qu'avec d'autres langages, [35].



Logo PYTHON

ANACONDA

Anaconda est une distribution libre et open source des langages de programmation Python et R appliquée au développement d'applications dédiées à la fouille de données et à l'apprentissage automatique (traitement de données à grande échelle, analyse prédictive, calcul scientifique), qui vise à simplifier la gestion des paquets et de déploiement. Les versions dépaquetages sont gérées par le système de gestion de paquets conda. La distribution Anaconda est utilisée par plus de 6 millions d'utilisateurs et comprend plus de 250 paquets populaires fouille de des données adaptés pour Windows, Linux et MacOS.

Anaconda est une distribution scientifique de Python. Permet d'écrire et d'exécuter le code Python de son choix par le biais du navigateur. Offert par anaconda Enterprise (gratuit), utilisé Jupyter Notebook et destiné à la formation et à la recherche dans l'apprentissage automatique, [36].



Interface ANACONDA

GOOGLE COLAB

- **Google Colab** : Colaboratory ou 'Colab'. Permet d'écrire et d'exécuter le code Python de son choix par le biais du navigateur. Offert par Google (gratuit), basé sur Jupyter Notebook et destiné à la formation et à la recherche dans l'apprentissage automatique. Cette plateforme permet d'entraîner des modèles de Machine Learning directement dans le Cloud. Colab permet :

- D'améliorer les compétences de codage en langage de programmation Python.
- De développer des applications en Deep Learning en utilisant des bibliothèques Python populaires telles que Keras, TensorFlow, PyTorch et OpenCV.
- D'utiliser un environnement de développement (Jupyter Notebook) qui ne nécessite aucune configuration, Mais la fonctionnalité qui distingue Colab des autres services est l'accès à un processeur graphique GPU, totalement gratuit, [37].



Interface Google Colab

1. **Files:** Vous pourrez télécharger des ensembles de données et d'autres fichiers depuis votre ordinateur et Google Drive.
2. **Extraits de code :** Vous pourrez trouver des extraits de code pré-écrits pour différentes fonctionnalités, telles que l'ajout de nouvelles bibliothèques ou le référencement d'une cellule à une autre.
3. **Exécuter la cellule :** il s'agit du bouton d'exécution. Cliquer dessus exécutera tout code inséré dans la cellule à côté. Vous pouvez utiliser le raccourci Maj + Entrée pour exécuter la cellule actuelle et quitter pour une nouvelle.
4. **Table des matières :** Vous pourrez créer et parcourir différentes sections à l'intérieur de votre bloc-notes. Les sections vous permettent d'organiser votre code et d'améliorer la lisibilité.
5. **Barre de menus :** comme dans toute autre application, cette barre de menus peut être utilisée pour manipuler le fichier entier ou ajouter de nouveaux fichiers. Parcourez les différents onglets et familiarisez-vous avec les différentes options. En particulier, assurez-vous de savoir comment télécharger ou ouvrir un bloc-notes et télécharger le bloc-notes (toutes ces options se trouvent sous "Fichier").
6. **Nom de fichier :** il s'agit du nom de votre fichier. Vous pouvez cliquer dessus pour changer le nom. Ne modifiez pas l'extension (.ipynb) lors de la modification du nom du fichier car cela pourrait rendre votre fichier impossible à ouvrir.
7. **Insérer une cellule de code :** ce bouton ajoutera une cellule de code sous la cellule que vous avez actuellement sélectionnée
8. **Insérer une cellule de texte :** ce bouton ajoutera une cellule de texte sous la cellule que vous avez actuellement sélectionnée.
9. **Cellule :** C'est la cellule. C'est ici que vous pouvez écrire votre code ou ajouter du texte en fonction du type de cellule dont il s'agit.
10. **Sortie :** il s'agit de la sortie de votre code, y compris les erreurs, qui seront affichées.

11. Effacer la sortie : ce bouton supprimera la sortie.

- **Avantages de Colab :** En plus d'être facile à utiliser, le Colab est assez flexible dans sa configuration et fait beaucoup de travail pour nous :
 - Prise en charge de Python 2.7 et Python 3.6
 - Accélération GPU gratuite
 - Bibliothèques préinstallées : les principales bibliothèques de Python comme TensorFlow, Scikit-learn, entre autres, sont préinstallées et prêtes à être importées.
 - Prend en charge les commandes Bash (langage de commande).

BIBLIOTHEQUES UTILISEES SONT :

1. NumPy

C'est l'une des bibliothèques les plus essentielles lorsqu'on travaille sur des projets en Python. Elle permet des manipulations de tableaux de grande dimension comme les opérations arithmétiques, les conversions, etc



2. OpenCV

OpenCV est une bibliothèque graphique libre, initialement développée par Intel, C'est l'une des meilleures bibliothèques de vision par ordinateur et de traitement d'image en temps réel utilisée pour faire des tonnes de choses sur des images et des vidéos.



3. Matplotlib

Matplotlib est une bibliothèque destinée à tracer et visualiser des données sous formes de graphiques, c'est-à-dire donner une représentation visuelle des résultats. Nous l'utilisons quand nous devons tracer quelques images pour faire des comparaisons. Elle peut être combinée avec les bibliothèques python de calcul scientifique NumPy et SciPy.



4. Pandas

Cette bibliothèque est utilisée lorsqu'on a besoin d'organiser nos données et de les rendre plus faciles à utiliser et à comprendre. Elle les organise dans des structures appelées « trames de données » et celles-ci aideront avec une représentation claire des données ainsi que leur personnalisation.



5. SciPy

C'est une bibliothèque Python open source pour le calcul scientifique et technique. Les principaux packages incluent des modules pour : l'optimisation, l'algèbre linéaire, l'intégration, la FFT, le traitement du signal et de l'image. SciPy utilise les tableaux et matrices du module NumPy.



6. Scikit-learn

Une bibliothèque Python open source qui propose de nombreuses classifications, des algorithmes de régression et de regroupement comme les support vector machines (SVM), le random forests, le gradient boosting, les k-means et DBSCAN. Elle est utilisée lors de la création et la formation des modèles de Machine Learning.



7. Tensorflow

TensorFlow est une bibliothèque d'apprentissage en profondeur open source développée par Google. Il s'agit d'une boîte à outils permettant de résoudre des problèmes mathématiques extrêmement complexes avec aisance.



8. Keras

Keras est un Framework Python pour l'apprentissage en profondeur. Son USP est la réutilisation du code pour le CPU et le GPU. Keras est compatible avec les versions de Python 2.7 / 3.6 et autres. Il a été initialement développé dans le cadre de l'effort de recherche du projet ONEIROS (système d'exploitation de robot intelligent neuro-électronique à composition open-ended). Keras peut désormais être exécuté sur TensorFlow.

