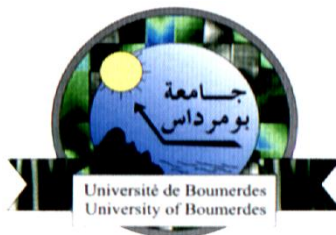


REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE
UNIVERSITE M'HAMED BOUGARA BOUMERDES



Faculté de Technologie
Département Génie Mécanique

Mémoire de Master

En vue de l'obtention du diplôme de **MASTER** en : Electromécanique

Filière : Electromécanique
Spécialité : Electromécanique

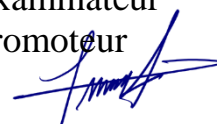
THEME

**Contribution à l'optimisation du nombre de capture en vue de
diagnostic intelligent des systèmes hydrauliques**

Présenté par :
ZEMOUR hicham
MALOUAJMI alaa dine

Soutenue publiquement le 15/07/2021, Devant le jury compose de :

IKHLEF	Boualem	MAA	President
GOUGAM	Fawzi	MAB	Examineur
RAHMOUNE	Chemseddine	MCA	Promoteur



Promotion 2020- 2021

N° Ordre...../DGM/FT/UMBB/2021

OK pour DSpace



Le 28/07/2021

Remerciement

*On remercie tout d'abord **Allah** le tout puissant de nous avoir donné la santé et la volonté de terminer ce modeste travail*

Ce travail ne serait pas aussi riche et n'aurait pas pu voir le jour, sans l'encadrement de Dr. RAHMOUNE, on le remercie pour son aide a bien diriger notre travail, sa patience et sa disponibilité durant la préparation de ce mémoire.

Aussi nous adressons nos vifs remerciements aux membres du jury pour l'intérêt qu'ils ont porté à notre recherche en acceptant d'examiner notre travail et de l'enrichir par leurs propositions, sans oublier de remercier tous ceux qui ont contribué de près ou de loin pour l'élaboration de ce travail...

ملخص:

نظرًا لأهمية النظام الهيدروليكي ووجوده في العديد من المجالات الصناعية ، فقد قمنا بدراسة وتطوير تقنيات تشخيص الأعطال الهيدروليكية لتلافي الخسائر.

لتصنيف. في عملنا ، اقترحنا تقنية فعالة تسمح لنا بتشخيص نظام هيدروليكي معقد بأقل عدد ممكن من أجهزة الاستشعار استخدمنا طريقة خوارزمية النمل لتحسين عدد أجهزة الاستشعار وتقنية الذكاء الاصطناعي KNN الاعطال

لإثبات فعالية الطريقة المقترحة ، قمنا بإجراء مقارنة النتائج التي تم الحصول عليها باستخدام طرق أخرى وفي النهاية اثبتنا فعالية ودقة التقنيه المستعمله

Résumé

compte tenu de l'importance des système hydraulique et de sa présence dans des nombreuses zones industrielles, nous avons étudié et développé des techniques de diagnostic des défaut hydrauliques pour éviter les pertes.

Dans notre travail, nous avons proposé une technique efficace qui nous permettrait de faire le diagnostic d'un système hydraulique complexe avec moins de nombre de capteurs possible.

Nous avons utilisé la méthode des algorithmes des fourmis pour optimisé le nombre de capteurs et la technique d'intelligence artificiel kNN pour la classification des défaut.

Pour démontré l'efficacité de la méthode proposée en a fait une comparaison des résultats obtenue par des autres méthodes.

Finalement, nous avons prouvé l'efficacité et la précision de la technique utilisée.

Liste des figure

Figure I.1 :	Structure schématique des systèmes hydrauliques	6
Figure I.2 :	Groupe générateur hydraulique	7
Figure I.3 :	Schéma d'une pompe centrifuge.	8
Figure I.4 :	pompe centrifuge	8
Figure I. 5 :	schémas d'une pompe volumétrique à engrenage	8
Figure I. 6 :	pompe volumétrique à engrenage	8
Figure I.7 :	distributeur hydraulique	8
Figure I.8 :	fonctionnement du distributeur	9
Figure I.9 :	accumulateur hydraulique	10
Figure I.10 :	Motors hydraulique	11
Figure I .11 :	vérin hydraulique	11
Figure I 12 :	Procédure de détection et d'isolation des défauts	13
Figure II.1 :	Architecteur général de l'apprentissage automatique	21
Figure II.2 :	schémas de déférent type d'apprentissage	22
Figure II.3	architecteur des réseaux de neurone	25
Figure II.4 :	Création de l'hyperplan de séparation optimal à l'aide de vecteur de support	26
Figure II.5 :	Un exemple de diagnostic d'un patient avec arbre de décision	27
Figure II. 6 :	Visualisation d'un exemple de classification en utilisant	28
Figure II.7 :	Organigramme de l'algorithme KNN	29
Figure III.1:	Procédure de sélection des caractéristiques	36
Figure III.2	La procédure du méthode "filtre"	39
Figure III.3 :	La procédure du La méthode Embedded (embarquée)	39
Figure III.4 :	La procédure du La méthode wrapper	40
Figure III.5 :	schéma de comportement des fourmis pour choisir le plus court chemin recherche de nourriture	41
Figure III.6 :	principe de l'algorithme des fourmis	43
Figure III.7 :	Le processus global de sélection d'attributs par algorithme des fourmis (ACO)	44
Figure IV.1	schéma de notre système hydraulique	46
Figure IV.2 :	organigramme de la méthode utilisé	49
Figure IV.3:	graphes des résultat accuracy	52
Figure IV.4 :	les matrice de confusion	53

Liste des tableau :

Tableau IV.1 :	résultat de classification par l' arbre de décision DT	54
Tableau IV.2:	résultats de la classification par KNN avec k=3	55
Tableau IV.3:	résultats de la classification par KNN avec k=5	56
Tableau IV. 4	résultats de la classification par KNN avec k=7	57
Tableau IV. 5:	résultats de la classification par KNN avec k=9	58

Sommaire

Introduction général	4
<u>Chapitre I : Généralité sur le diagnostic des systèmes hydraulique</u>	
I.1. Introduction.....	5
I .2. Définition d'un système Hydraulique	5
I.3. Domaine d'application des systèmes hydrauliques	5
I.4. Les composant d'un système hydraulique.....	6
I.4.1 Le groupe générateur.....	7
I.4.2 Les pompe	7
I.4.2.1. Les pompe centrifuges.....	7
I.4.2.2. Les pompes volumétriques	8
I.4.3 La partie distribution.....	8
I.4.4. Le contrôle	9
I.4.4.1. Limiteur de pression	9
I.4.4.2. Le réducteur de pression à action direct	9
I.4.4.3. Limiteur de débit	9
I.4.4.3.1. La valve :	10
I.4.4.3.2. Les valves de séquence :	10
I.4.4.3.3. La valve d'équilibre :	10
I.4.4.4. Les Accumulateurs :.....	10
I.4.5.1. Définition d'un actionneur	11
I.4.5.2. Les moteurs hydrauliques.....	11
I.4.5.3. Les vérins hydrauliques :.....	11
I.5. Généralité sur le Diagnostic	12
I.5.1. Définition de quelque terme sur le diagnostic	12
I.5. 2. Procédure de détection et d'isolation des défauts	13
I.5.3. Les méthodes de diagnostic	14
I.5.3.1 Les méthodes basées sur un modèle.....	14
I.5.3.2 Les méthodes basées sur l'intelligence artificielle	14
I.5.4. LA REDONDANCE :	15
I.5.4.2 La redondance analytique	15
I.6. Les Techniques de diagnostic des systèmes hydraulique	15
I.6. 1. L'analyse vibratoire	16

I.6. 2. L'analyse par thermographie infrarouge	16
I.6.2.1. Principe de l'analyse thermographique.....	16
I.6.3. Analyse acoustique ultrasonore	16
I.6.3.1 Principe du contrôle par ultrasons	16
I.6.4. L'analyse des huiles	17
I.6.4.1 La mise en place des méthodes d'analyse des huiles	17
I.7. Conclusion	18
<u>Chapitre II : Généralité sur l'intelligence artificielle et l'apprentissage automatique</u>	
II.1 Introduction	19
II.2. Intelligence Artificielle	19
II.3. Les applications de l'intelligence artificielle	20
II.4. Apprentissage automatique	20
II.4.1 Les types d'apprentissages Automatiques	21
II.4.1.1 L'apprentissage supervisé	22
II.4.1.1.1 Régression	22
II.4.1.1.2. Classification supervisée.....	23
I.4.2 Les méthodes de Classification supervisée.....	24
I.4.2.1 Les réseaux de neurones artificiels	24
I.4.2.2 La machine à support vectoriel (SVM)	25
I.4.2.3. Arbre de décision.....	26
I.4.2.4 Forêts aléatoires.....	27
I.4.2.5. Méthode des K plus proches voisins	27
I.4.3. L'évaluation de performance des modèle d'apprentissage	29
I.4.3.1 Taux de classification (accuracy) et Taux d'erreur (erreur accuracy)	29
I.4.3.2. Matrice de confusion	30
I.5. Conclusion	30
<u>Chapitre III : Notion d'optimisation et de la sélection des caractéristiques</u>	
II.1. Introduction	31
III.2. Quelques notions sur l'optimisation.....	31
III.2.1. Définition d'Optimisation	31
III.2.2. Formule de l'optimisation	31
III.3. Les domaines d'application de l'optimisation.....	32
III.4. Notion sur la sélection des caractéristiques.....	33
III.4.1 Définition de la sélection des caractéristiques	33

III.4.2 Objectifs de la sélection des caractéristiques	33
III.4.3 La sélection vue comme un problème d'optimisation	34
III.4.4. Procédure de la sélection des caractéristiques	35
III.4.4.1. Les procédures de recherches	36
III.4.4.1.1. Génération exhaustive.....	36
III.4.4.1.2. Génération heuristique	36
III.4.4.1.3. Génération aléatoire.....	37
III.4.4.2. La fonction d'évaluation	37
III.4.4.3. Validation du sous ensemble	38
III.4.5. Les méthodes de sélection d'attributs	38
III.4.5.1 Les méthodes filtres.....	38
III.4.5.2 La méthode Embedded (embarquée)	39
III.4.5.3. Wrapper méthode	39
III.5. L'algorithme d'optimisation (aco) pour la sélection des	40
Caractéristiques.....	40
III.5.1 Le comportement des fourmis pour cherche la nourriture	40
III.5.2 Le principe de l'algorithme des fourmis (ant colony optimisation)	41
III.5. Conclusion	44

Chapitre IV : application d'algorithme de fourmis et KNN pour la sélection et la classification des caractéristiques

IV.1. Introduction.....	45
IV.2. Description de système et de base des donné utilisé	45
IV.3. Conception et la méthodologie du la technique utilisé	48
IV.3.1 Les défaut traité dons notre travail	48
IV.3.2. Extraction des caractéristiques	50
IV.3.3. La sélection et la classification des caractéristiques.....	50
IV.3.4. Evaluation et interprétation des résultats obtenue	51
IV..3.4.1. Les résultats de la sélection.....	51
IV..3.4.2. Résultat de la classification.....	52
IV.4. Conclusion.....	58
Conclusion général	59
Bibliographié.....	60

Introduction générale :

L'hydraulique est considéré comme une révolution industrielle en raison de l'utilisation des systèmes hydraulique dans des nombreux domaines industriels, Les systèmes hydrauliques sont largement utilisés dans des domaines importants comme l'aéronautique, l'agriculture, la sidérurgie, la production des énergies etc. Ces systèmes sont parfaitement adaptés pour transporter de l'énergie ou pour piloter des actionneurs (vérins, moteur, etc.), pour le transport des fluide (l'eau, le pétrole, les huile ...) la lubrification et le refroidissement de diverse machine.

Les défaillances du système hydraulique causées par des conditions sévères et complexes peuvent entraîner des accidents catastrophiques, d'énormes pertes économiques financières, commerciales et environnementales.

Donc le développement d'une technique efficace pour diagnostiquer les système hydrauliques est une tâche essentielle.

Pour ces raisons notre contribution dans ce mémoire est de proposer une technique efficace et stable pour diagnostiquer les systèmes hydrauliques.

I.1. Introduction

Les systèmes hydrauliques sont largement utilisés dans des domaines importants comme l'aéronautique, l'agriculture, la sidérurgie, la production des énergies etc. Ces systèmes sont parfaitement adaptés pour transporter de l'énergie ou pour piloter des actionneurs (vérins, moteur, etc.), Pour le transport des fluide (l'eau, le pétrole, les huile ...) la lubrification et le refroidissement de divers machine. Grâce à cette importance en a besoin de maintenir ces systèmes à un bon fonctionnement pour éviter les défailant ces qui provoquent des dégâts, qui minimisent le rendement et qui fant des pertes de production, D'où le diagnostic des défauts son premier degré sera nécessaire pour maintenir ces systèmes.

Dans ce chapitre, nous parlerons des systèmes hydrauliques en général et des moyens de les diagnostiquer.

I .2. Définition d'un système Hydraulique

Un système hydraulique est un assemblage de composant qui fonctionnant de manière unitaire pour effectuer un travail mécanique (entraînent les parties mobiles de nombreux équipements industriels) en utilisant la force transmise par un fluide sous pression.

En d'autres mots, c'est l'utilisation des fluides sous pression et des lois des liquides pour permettre un travail mécanique.

Les systèmes hydrauliques utilisent des fluides non compressibles, Ces liquides sont propulsés dans le système grâce à la pompe et dirigés selon l'action désirée.

I.3. Domaine d'application des systèmes hydrauliques

On trouve les systèmes hydrauliques dans nombreux machine (Les machine-outil, les engins de travaux publics, les avions, les véhicules industriels, les machines agricoles Les chariots élévateurs...).

Nous nous concentrons sur elle en raison de sa présence dans de nombreux domaines de grande importance, tels que le secteur de la production d'énergie) stations d'extraction de pétrole brut, les centrales de la production d'électricité), où il est utilisé pour :

- Le Transport de gaz et divers fluide (l'eau, les huile, le pétrole.)
- Le refroidissement des diverses machines (les turbine, les alternateur...)

-La lubrification des nombreuses machines

I.4. Les composant d'un système hydraulique

Le système d'entraînement hydraulique se compose en règle générale d'une unité d'alimentation en pression dotée d'une pompe hydraulique et actionneur qui transforme l'énergie hydraulique transmise par la pompe en énergie mécanique (les vérin). L'entraînement est commandé par le biais d'éléments de commande, par exemple des vannes.

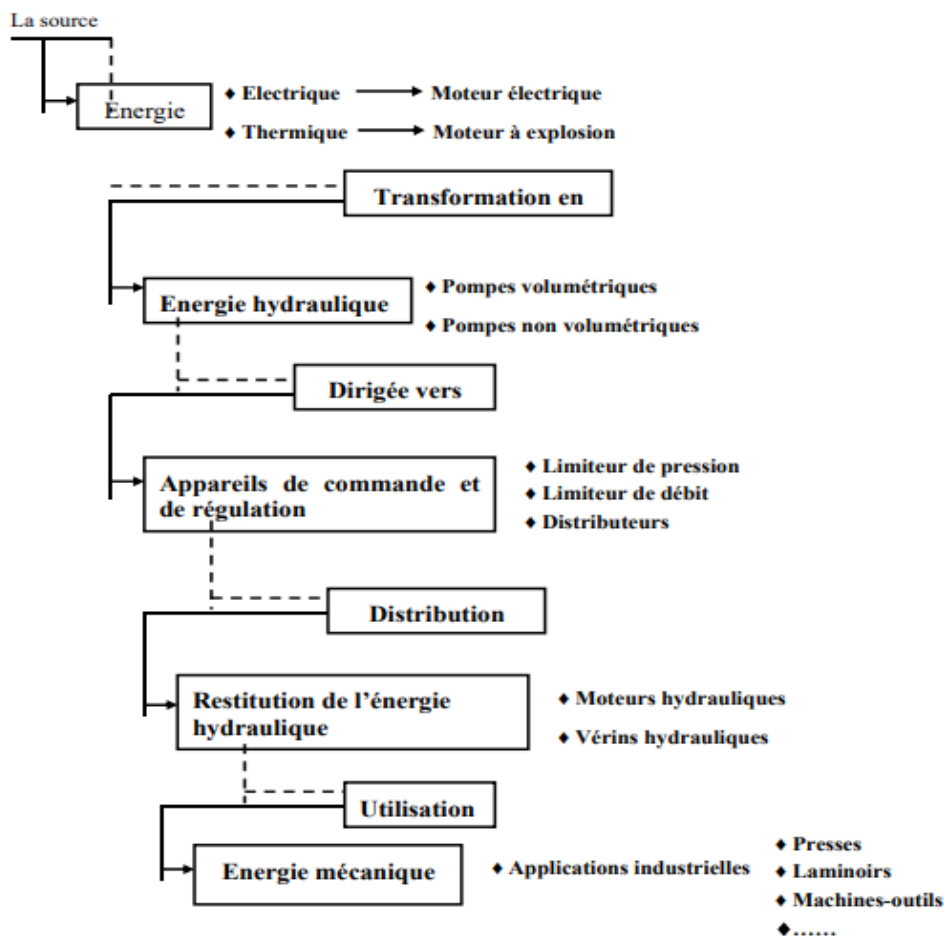


Figure I.1 : Structure schématique des systèmes hydrauliques

Plusieurs éléments composent les différents systèmes hydrauliques en peut les diviser en fonction de leur rôle en quatre parties : le groupe générateur, la partie distribution, le contrôle et les actionneurs.

I.4.1 Le groupe générateur

Permet la circulation du fluide dans le système hydraulique. Il se compose généralement d'une pompe, d'un moteur d'entraînement et d'un réservoir. Le liquide est mis sous pression par la pompe pour ensuite créer un débit. Lorsque l'opération est terminée, le fluide se dirige vers le réservoir.



Figure I.2 : Groupe générateur hydraulique

I.4.2 Les pompe

Une pompe est définie comme étant une machine capable de déplacer un liquide d'un point A à un point B en transformant l'énergie mécanique qu'elle reçoit en énergie hydraulique. Diverses technologies de pompe sont apparues au fil du temps en fonction du besoin et des conditions d'utilisation. Néanmoins, deux grandes familles de pompes peuvent être distinguées : les pompes roto dynamiques et les pompes volumétriques. [1]

I.4.2.1. Les pompe centrifuges

Une pompe centrifuge est le type de pompe le plus commun c'est une machine rotative qui pompe un liquide en le forçant au travers d'une roue à aube ou d'une hélice appelée impulseur, Le fluide arrivant par l'ouïe est dirigé vers la roue en rotation qui sous l'effet de la force centrifuge lui communique de l'énergie cinétique. Cette énergie cinétique est transformée en énergie de pression dans la volute. Un diffuseur à la périphérie de la roue permet d'optimiser le flux sortant est ainsi de limiter les pertes d'énergie. [2]

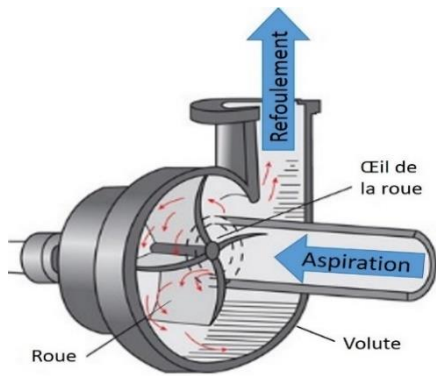


Figure I.3 : Schéma d'une pompe centrifuge.



Figure I.4 : pompe centrifuge

I.4.2.2. Les pompes volumétriques

Une pompe volumétrique se compose d'un corps de pompe parfaitement clos à l'intérieur duquel se déplace un élément mobile rigoureusement ajusté. Ce déplacement est cyclique. Pendant un cycle, un volume de liquide pénètre dans un compartiment avant d'être refoulé.

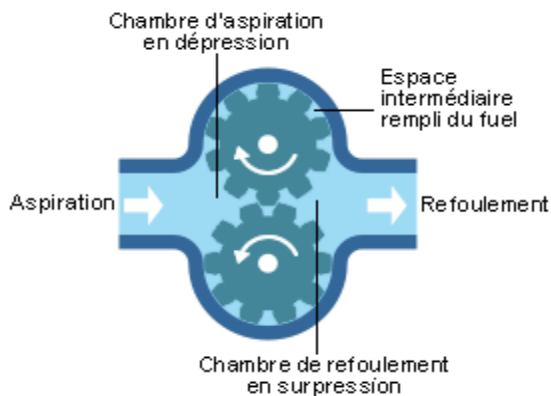


Figure I. 5 : schémas d'une pompe volumétrique à engrenage



Figure I. 6 : pompe volumétrique à engrenage

I.4.3 La partie distribution

Les distributeurs appelés sont des pré actionneurs qui orientent la circulation du fluide dans diverses directions, assurent l'alimentation des actionneurs et les retours de fluide à la bêche.

[3]



Figure I.7: distributeur hydraulique

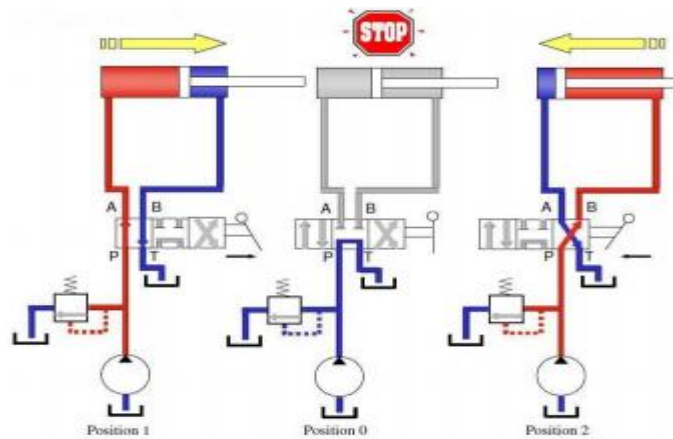


Figure I.8 : fonctionnement du distributeur

En Rouge l'huile sous pression

- En bleu l'huile sans pression (circuit de retour)
- Sens de déplacement du vérin dans chaque position.

I.4.4. Le contrôle

Assure la gestion de la pression et du débit du système. Selon le travail à effectuer, on peut appliquer plus ou moins de pression.

I.4.4.1. Limiteur de pression

Il protège le système de trop hautes pressions. Si la pression requise pour soulever la charge est trop élevée, ce limiteur s'ouvre et relâche la pression en renvoyant le liquide dans le réservoir. L'utilisation du limiteur de pression est également nécessaire lorsque le piston atteint la fin de sa course. À ce moment-là, le liquide n'a pas d'autre choix que de retourner dans le réservoir par le limiteur de pression.

I.4.4.2. Le réducteur de pression à action direct

Il est normalement ouvert au repos et lorsque la pression en aval dépasse la valeur de tarage, elle coupe la communication entre l'entrée et la sortie. [4]

I.4.4.3. Limiteur de débit

Le contrôle de débit Le rôle de ce composant est de faire varier la section dans laquelle le fluide circule. [4]

I.4.4.3.1. La valve :

Nom générique d'un appareil avec une partie mobile lui permettant d'ouvrir et de fermer une voie de passage afin de permettre, d'empêcher ou de réguler le flux d'un fluide.

I.4.4.3.2. Les valves de séquence :

Elle permet d'alimenter un circuit secondaire lorsqu'une certaine pression est atteinte dans le circuit primaire. [5]

I.4.4.3.3. La valve d'équilibre :

Son rôle est de maintenir en position un récepteur hydraulique (vérin ou moteur). Elle peut être envisagée avec deux pilotages.

I.4.4.4. Les Accumulateurs :

Leur rôle est de stocker un certain volume de fluide sous pression pour le restituer en fonction des besoins. Ils sont aussi utilisés pour les commandes d'urgence (terminer un mouvement, actionner un frein,) ou pour amortir les chocs dans les démarrages des installations. Le principe est simple : une chambre à deux orifices séparés par un élément étanche. L'un des orifices est relié au système et l'autre permet le remplissage avant utilisation d'un gaz générant un contre effort. Le gaz utilisé est l'azote (gaz inerte). [5]



Figure I.9 : accumulateur hydraulique

I.4.5. Les actionneurs

I.4.5.1. Définition d'un actionneur

Consiste à transformer l'énergie hydraulique transmise par le fluide en énergie mécanique. Dans un circuit, l'actionneur hydraulique constitue l'outil indispensable pour convertir l'énergie hydraulique en énergie mécanique grâce à un fluide sous pression

I.4.5.2. Les moteurs hydrauliques

Ils fournissent une transmission d'énergie inverse de celle des pompes. Ils convertissent l'énergie hydraulique en énergie mécanique. Dans leurs conceptions, ils sont très proches des pompes



Figure I.10 : Motors hydraulique

I.4.5.3. Les vérins hydrauliques :

Les vérins sont des moteurs (ou pompes) linéaires. On considère généralement que les fuites internes de ces organes sont négligeables, et donc que leur rendement volumétrique est proche



Figure I .11 : vérin hydraulique

Chapitre I : Généralité sur le diagnostic des systèmes hydraulique

De 1 On ne décrira ici que les vérins à tige simple, le raisonnement pour les vérins à double tige est analogue. Ils sont constitués d'un corps fixe et d'un ensemble piston tige mobile (ou l'inverse). Des variantes permettent d'obtenir des freinages en fin de course. [6]

I.5. Généralité sur le Diagnostic

I.5.1. Définition de quelque terme sur le diagnostic

Le diagnostic de défauts est une procédure qui consiste à détecter, localiser et identifier un composant ou un élément défectueux dans un système dynamique. [8] :

Le système de diagnostic permet la détection précoce de déviations faibles par rapport à une caractérisation du système en fonctionnement nominal, afin de prévenir un dysfonctionnement avant qu'il n'arrive. [8]

Pour mieux étayer nos propos, il est important de définir quelques termes associés à la supervision des procédés :

Un Fonctionnement normal d'un système : Un système est dit dans un état de fonctionnement normal lorsque les paramètres qui le caractérisant demeurent au voisinage de leurs valeurs nominales. Le système est dit défaillant dans le cas contraire. [8]

Un défaut : L'altération ou la cessation de l'aptitude d'un ensemble à accomplir ses fonctions requises avec des performances définies dans les spécifications techniques. [7]

Une défaillance : est la cause d'une anomalie, il implique l'apparition d'un défaut car il y a un écart par rapport au caractéristique théorique, par contre un défaut n'implique pas nécessairement une défaillance puisque le dispositif peut très bien accomplir sa fonction principale.

Une dégradation d'un procédé : c'est le processus qui amène à un état défaillant du système.

Une panne : Etat d'un système incapable d'assurer le service spécifié à la suite d'une défaillance [7]

Modèles des défauts : Un modèle de défaut est une représentation formelle de la connaissance des défauts et de leurs façons d'influencer le système. En général, l'utilisation de modèles de défauts assure une meilleure performance du diagnostic. Plus précisément, des défauts de faibles amplitudes peuvent être détectés et différents types de défauts peuvent être localisés [10].

Un diagnostic sur système consiste à faire une comparaison entre l'information mesurée, lors du fonctionnement réel du système, et la connaissance a priori de ses modes de fonctionnement. En se basant sur les résultats obtenus de cette comparaison, l'utilisateur peut intervenir et mettre en place les actions correctrices nécessaires à un retour à la normale.

Le concept de redondance analytique repose sur l'utilisation d'un modèle mathématique du système à surveiller [9]

I.5. 2. Procédure de détection et d'isolation des défauts

On trouve dans le diagnostic deux tâches essentielles sont l'observation des symptômes de défaillance ainsi que l'identification de la cause de défaillance à l'aide d'un raisonnement logique fondé sur des observations acquises sur le système, à partir des informations disponibles sur son état de fonctionnement, l'objectif est de détecter, localiser et identifier les défaillances qui peuvent affecter sa sûreté de fonctionnement.

Le diagnostic est constitué, dans la plupart des cas, de plusieurs étapes, figure I 12 :

La détection : qui consiste à utiliser un jeu de mesures (indicateurs de défauts) qui permettent d'engendrer des symptômes. On parle aussi d'attributs quand on fait référence à l'utilisation de méthodes de classification pour la détection.

La localisation : consiste à déterminer de manière plus approfondie les composants défaillants ou la cause probable de la défaillance.

L'identification des défauts : a pour but de déterminer le type (classe, taille, ...) d'un défaut (classification).

La prise de décision : doit permettre de générer, éventuellement sous le contrôle d'un opérateur humain, les actions correctrices nécessaires à un retour au fonctionnement normal de l'installation.

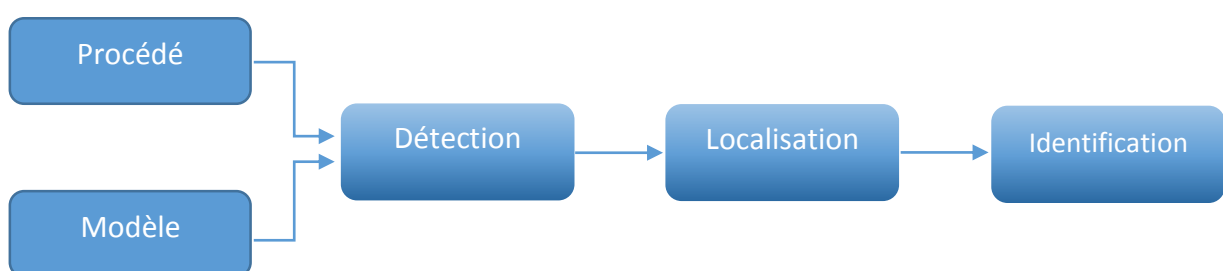


Figure I 12 : Procédure de détection et d'isolation des défauts

I.5.3. Les méthodes de diagnostic

On peut globalement distinguer deux grandes familles dans les méthodes de diagnostic :

- Les méthodes basées sur une modélisation des systèmes ou des signaux.
- Les méthodes basées sur l'intelligence artificielle

I.5.3.1 Les méthodes basées sur un modèle

Les méthodes de diagnostic à base de modèle peuvent être divisées en deux parties : méthodes à base de 'modèle mathématique du système' et méthodes à base de 'modèle du signal' [10].

D'un côté, les méthodes à base de modèle mathématique visent à représenter la dynamique du système par des équations mathématiques qui sont ensuite exploitées pour le diagnostic la modélisation des systèmes dynamiques souvent formulée par des relations non linéaires qui compliquent l'estimation des grandeurs d'intérêt. De l'autre côté, les méthodes à base de modèle du signal visent à modéliser le signal mesuré et à exploiter ses paramètres (phase, l'amplitude, moyenne, etc.) pour faire le diagnostic.

Le principe du diagnostic à base de modèle mathématique est de comparer le comportement actuel du système avec son comportement attendu (état sain). Cette comparaison génère des quantités résiduelles qui servent à prendre la décision du diagnostic. [10]

I.5.3.2 Les méthodes basées sur l'intelligence artificielle

Les méthodes basées sur l'intelligence artificiel fonctionnent sur le même principe que celles basées sur des modèles quantitatifs, mais la relation entre les descripteurs est exprimée par des fonctions qualitatives. Ces méthodes ne sont pas basées sur l'écriture de lois physiques

Le diagnostic avec l'intelligence artificielle repose sur l'utilisation des algorithmes d'apprentissage automatique qui sont soumis à deux étapes :

-L'étape de l'apprentissage : Dans cette phase (entraînement ou apprentissage), l'objectif est de trouver, à partir de l'ensemble de données expérimentales, les caractéristiques du comportement du système qui vont permettre de différencier les états dans lesquels peut se trouver le système. Cela concerne la conception et la construction d'un système de

Chapitre I : Généralité sur le diagnostic des systèmes hydraulique

classification, Il se compose de deux phases principales : une phase d'apprentissage (supervisé ou non supervisé) et l'attribution des classes obtenues dans d'importants états fonctionnels. [8]

Phase de reconnaissance :

Lors de la deuxième phase, la reconnaissance *de* nouvelles données est effectuée. Ce qui permettra d'identifier l'état fonctionnel de la procédure à partir des données en ligne des capteurs et d'actionneurs ou d'autres informations à partir des variables de processus. Les données sont classées en fonction du classifieur obtenu lors de la phase précédente [8]

I.5.4. LA REDONDANCE :

I.5.4.1 Redondances matérielle

Le principe de base du diagnostic des défauts repose sur la notion de redondance, qui fournit au système plusieurs informations différentes sur une même variable. Des tests vont alors permettre de vérifier la cohérence de ces informations. [7]

I.5.4.2 La redondance analytique

Un complément à la redondance physique consiste à exploiter les contraintes liant les différentes variables du système. Ces contraintes peuvent souvent s'exprimer sous la forme de relations analytiques liant les variables connues (relations d'entrée/sortie ou de sortie/sortie). Ces relations sont appelées relations de redondances analytiques.

I.6. Les Techniques de diagnostic des systèmes hydraulique

La diagnostic d'un équipement de machine est assurée en relevant périodiquement un indicateur d'état de dégradation ou de performance, il existe différentes techniques d'analyse tels que :

- 1 -la mesure de vibrations.
- 2 -la thermographie infrarouge.
- 3 -l'analyse des mesures ultrasonores.
- 4 -l'analyse d'huile.

I.6. 1. L'analyse vibratoire

L'analyse est une technique de contrôle non destructif employée pour réaliser, analyser l'état des installations industrielles dans le but d'opérer la maintenance préventive conditionnelle par surveillance. On pourra par exemple optimiser la conception en supprimant les fréquences de résonance qui provoquent les déformations de structure, détecter et identifier les défaillances d'un système. Une analyse vibratoire ponctuelle permet également d'émettre un diagnostic sur l'état général des composants mécaniques d'un équipement. La surveillance vibratoire est particulièrement adaptée aux machines tournantes (boîte de vitesse, réducteur, pompe, compresseur, ventilateur, accouplement, palier, roulements, denture, poulie).

I.6. 2. L'analyse par thermographie infrarouge

La thermographie infrarouge « TIR » est la science de l'acquisition et de l'analyse d'informations thermiques à l'aide de dispositifs d'imagerie thermique à distance

I.6.2.1. Principe de l'analyse thermographique

Tout corps dont la température est supérieure à zéro degré absolu émet un rayonnement électromagnétique. Et on est capable de détecter ce rayonnement sous la forme de sensation de chaleur. La thermographie infrarouge est la technique permettant de mesurer, par l'intermédiaire d'un détecteur, la puissance du rayonnement électromagnétique dans le spectre des infrarouges, émis par chaque point d'une scène ou d'un objet observé. Le détecteur de la caméra infrarouge reçoit un rayonnement, le convertit en signal électrique et reconstitue sur un écran une image thermique visible de l'objet émetteur. Cette image est la thermographie. Le thermo gramme est constitué par l'ensemble des points de valeurs de mesures thermiques. Ces valeurs thermiques en (°C) sont obtenues par inscription des valeurs radiométriques données par la caméra infrarouge [11]

I.6.3. Analyse acoustique ultrasonore

La technique d'analyse acoustique offre de multiples applications dans le contrôle d'équipements hydrauliques et pneumatiques (détection de fuites, contrôle d'étanchéité)

I.6.3.1 Principe du contrôle par ultrasons

Chapitre I : Généralité sur le diagnostic des systèmes hydraulique

Le contrôle par ultrasons comprend 3 étapes :

- La vibration mécanique engendrée par le palpeur émetteur se propage dans la pièce en se réfléchissant aux interfaces.
- Une partie du faisceau est interceptée par le défaut et renvoyée vers le palpeur récepteur (Eventuellement le même que l'émetteur) qui convertit la vibration en signal électrique.

On observe sur l'écran de visualisation un écho caractéristique apparaissant à une distance donnée sur la base de temps. [9]

I.6.4. L'analyse des huiles

L'analyse d'huile est une des méthodes les plus intéressantes pour assurer la santé et la longévité des machines. Sa mise en œuvre est simple et les informations fournies permettent des diagnostics précis. De plus, l'investissement est faible en contrepartie de l'information apportée et des coûts générés par des arrêts non programmés.

I.6.4.1 La mise en place des méthodes d'analyse des huiles

Sur site industriel

Par prélèvement d'échantillons et examens visuels (transparence, couleur, dépôts) des lubrifiants en service. Par un suivi continu de l'évolution des paramètres techniques tels que température, débit, pression de fonctionnement et les historiques des vidanges d'huile.

En laboratoire :

Par analyses physico-chimiques évaluant la qualité lubrifiante de l'huile, par la détermination de la teneur en produits d'usure, par examen microscopique et comptage de particules en suspension dans l'huile, la maintenance conditionnelle par l'analyse des huiles représentera un outil de progrès à la disposition des responsables de service maintenance. [9]

Exemple de détection et de suivi des dysfonctionnements industriels :

Chapitre I : Généralité sur le diagnostic des systèmes hydraulique

L'analyse d'huile, pratiquée dans le cadre d'une maintenance préventive conditionnelle, va permettre, par exemple, de détecter et de suivre les dysfonctionnements potentiels suivants :

Sur systèmes hydrauliques :

- Pollution du circuit (matières solides eau...).
- Usure des composants (pompes, moteurs, distributeurs...).
- Efficacité de la filtration (niveau de pollution...).
- Caractéristiques résiduelles du lubrifiant.

I.7. Conclusion

Dans ce chapitre nous avons présenté des généralités sur les systèmes hydrauliques et quelques notions sur le diagnostic et leur principe. Nous avons également discuté sur les diverses techniques de diagnostic utilisé pour le diagnostic des systèmes hydrauliques la plus par de ces méthodes est traditionnel comme (la mesure de vibrations, L'analyse des mesures ultrasonores, et l'analyse d'huile.) qui sont basées sur une modélisation des systèmes ou des signaux, il existe de nouvelles méthodes modernes basé sur l'intelligence artificiel et l'apprentissage automatique on a utilisé dans notre projet. Nous le mentionnerons dans le prochain chapitre les notion de notre travail.

Chapitre II : Généralité sur l'intelligence artificielle et l'apprentissage automatique

II.1 Introduction

Au milieu du vingtième siècle, le monde a été témoin d'un développement de l'informatique, et en conjonction avec les découvertes en neurosciences et le développement du contrôle automatisé, les chercheurs ont commencé à réfléchir à la création et au développement de dispositifs qui simulent la pensée humaine en simulant le système neuro. Un nouveau terme a été inventé : science de l'ordinateur, connu sous le nom d'intelligence artificielle. La première conférence sur l'intelligence artificielle s'est tenue à la faculté de Dartmouth en été 1956.

Depuis lors jusqu'à nos jours, l'intelligence artificielle a connu un développement terrible, et l'homme l'a mise dans de nombreux domaines, tels que le diagnostic médical et le diagnostic des défauts des machines industrielles.

Dans ce chapitre nous définirons l'intelligence artificielle et ses applications et mentionnons une de ses branches (apprentissage automatique) utilisée dans le diagnostic.

II.2. Intelligence Artificielle

L'intelligence Artificielle se définirait comme étant « l'ensemble de théories et de techniques mises en œuvre en vue de réaliser des machines capables de simuler l'intelligence. » Ce serait, de ce fait, des ordinateurs ou des machines dotées de programmes capables de performances similaires à l'intelligence humaine, ou même, amplifiées par la technologie. Ces machines sont en mesure de :

- Capacité de raisonner
- Capacité de traiter de grandes quantités de données
- La capacité de distinguer des modèles que les humains ne peuvent pas détecter.
- Capacités à interagir avec l'homme.
- Capacité d'apprendre progressivement. [12]

L'intelligence artificielle a besoin de deux choses de base : une grande collection de données et des algorithmes qui traitent ces données.

Chapitre II : Généralité sur l'intelligence artificielle et l'apprentissage automatique

II.3. Les applications de l'intelligence artificielle

Il existe de nombreux domaines d'application de l'IA, souvent avec d'autres disciplines informatiques.

- Robotique : construction de plans et suivi d'exécution, décision de la route à suivre (milieu hostile ou mal connu).
- le secteur des transport (les véhicules autonomes).
- Systèmes de diagnostic (médical, industrielle).
- Développement de vision où la machine peut capturer des informations visuelles, puis les analyser.
- L'apprentissage automatique (Une méthode de déterminer l'objectif puis le moyen de l'atteindre).

II.4. Apprentissage automatique

L'apprentissage automatique est un domaine sort de l'intelligence artificiel s'intéressant à comprendre et reproduire la faculté de l'apprentissage humain par des systèmes artificiels. Il s'agit, très schématiquement, de concevoir des algorithmes et des méthodes permettant d'extraire l'information pertinente de données, ou d'apprendre des comportements à partir d'exemples. Ainsi, le but essentiel de l'apprentissage automatique est de déterminer la relation (modèle) entre les objets et leurs catégories pour la prédiction et la découverte des connaissances [13].

Les données utilisées par l'apprentissage automatique donne le processus de modélisation sont appelées données de (formation), en anglais Training Data.

Chapitre II : Généralité sur l'intelligence artificielle et l'apprentissage automatique

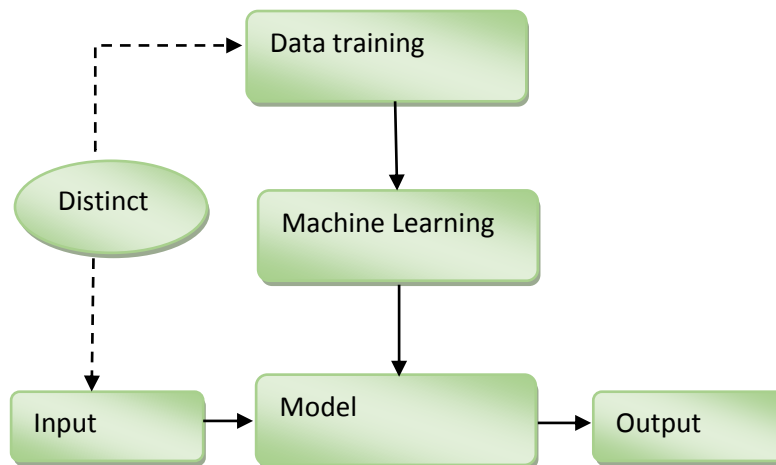


Figure II.1 : Architecteur général de l'apprentissage automatique

Training data : donnée de l'entraînement

Machine learning : algorithme capable d'apprendre les données

Model : modèle obtenu après l'apprentissage

Input : Les nouvelles entrées après le processus d'apprentissage et la préparation du système Intelligent pour le travail (distinct de la base de données).

Output : la sortie obtenue après l'injection de nouvelles entrées (input).

II.4.1 Les types d'apprentissages Automatiques

1)- **Supervise :** on dispose d'un ensemble d'objets et pour chaque objet une valeur cible associée ; il faut apprendre un modèle capable de prédire la bonne valeur cible d'un objet nouveau.

2)- **Non supervise :** on dispose d'un ensemble d'objets sans aucune valeur cible associée ; il faut apprendre un modèle capable d'extraire les régularités présentes au sein des objets pour mieux visualiser ou appréhender la structure de l'ensemble des données

3)- **Par renforcement :** on dispose d'un ensemble de séquences de décisions (politiques ou stratégiques) dans un environnement dynamique, et pour chaque action de chaque séquence une valeur de récompense (la valeur de récompense de la séquence est alors la somme des valeurs des récompenses des actions qu'elle met en œuvre) ; il faut apprendre un modèle capable de prédire la meilleure décision à prendre étant donné un état de l'environnement. [14]

Chapitre II : Généralité sur l'intelligence artificielle et l'apprentissage automatique

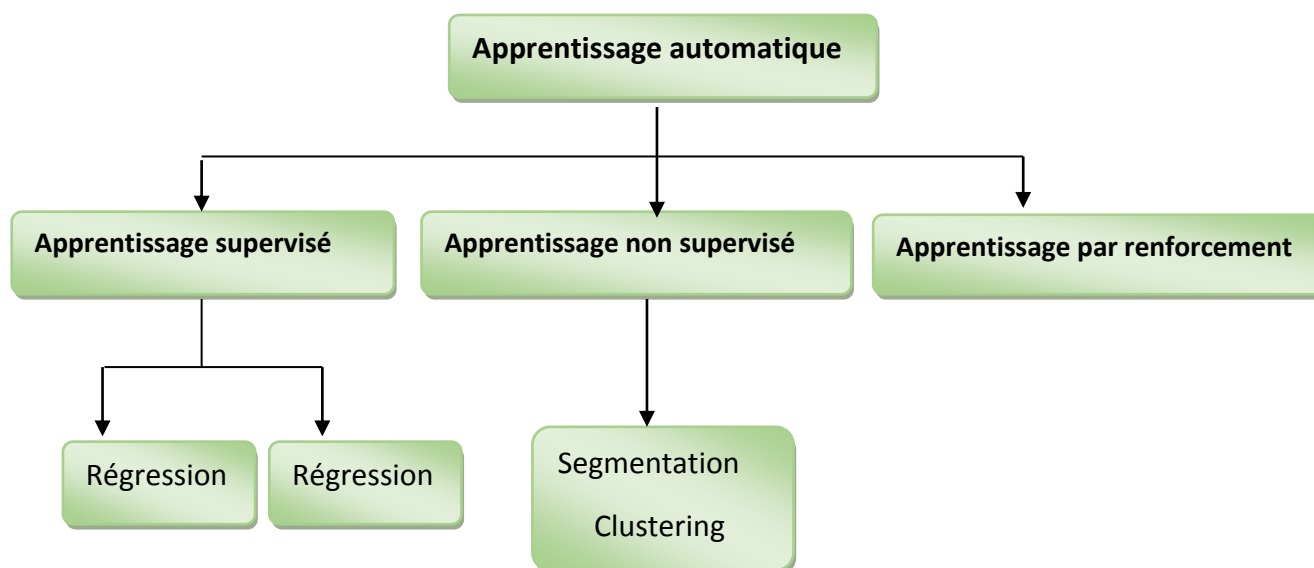


Figure II.2 : schémas de différents types d'apprentissage

II.4.1.1 L'apprentissage supervisé

L'algorithme est entraîné en utilisant une base de données d'apprentissage contenant des exemples de cas réels traités et validés. L'objectif est de trouver des corrélations entre les données d'entrée (variables explicatives) et les données de sorties (variables à prédire), pour ensuite inférer la connaissance extraite sur des entrées avec des sorties inconnues.

Chaque exemple, appelé aussi instance, est un couple d'entrée-sortie (x_n, y_n)

$1 \leq n \leq N$, avec $x_n \in X$ et $y_n \in Y$ et où :

X est l'ensemble d'attributs (discrets ou continus),

Y est l'ensemble des valeurs de sortie (la variable cible ou dépendante), Y peut être discrète ou continue [15].

En apprentissage supervisé, on distingue entre deux types de tâches (régression, classification)

II.4.1.1.1 Régression

Quand la variable cible à prédire est continue $Y \in \mathbb{R}$. [16]

Chapitre II : Généralité sur l'intelligence artificielle et l'apprentissage automatique

Dans l'apprentissage automatique, le but de la régression est prédire une valeur de sortie à partir des valeurs d'un ensemble de caractéristiques en entrée. Par exemple, prédire le prix d'une maison en se basant sur sa surface, nombre des étages, son emplacement, etc.

II.4.1.1.2. Classification supervisée

L'objectif de la classification supervisée est principalement de définir des règles permettant de classer des objets dans des classes à partir de variables qualitatives ou quantitatives caractérisant ces objets). On dispose au départ d'un échantillon d'apprentissage (Données d'entraînement) dont le classement est connu. Cet échantillon est utilisé pour l'apprentissage des règles de classement (modèle de classification).

Il est nécessaire d'étudier la fiabilité de modèle pour les comparer et les appliquer, évaluer les cas de sous apprentissage ou de sur apprentissage (complexité du modèle). On utilise souvent une nouvelles données échantillon indépendant pour teste le modèle.

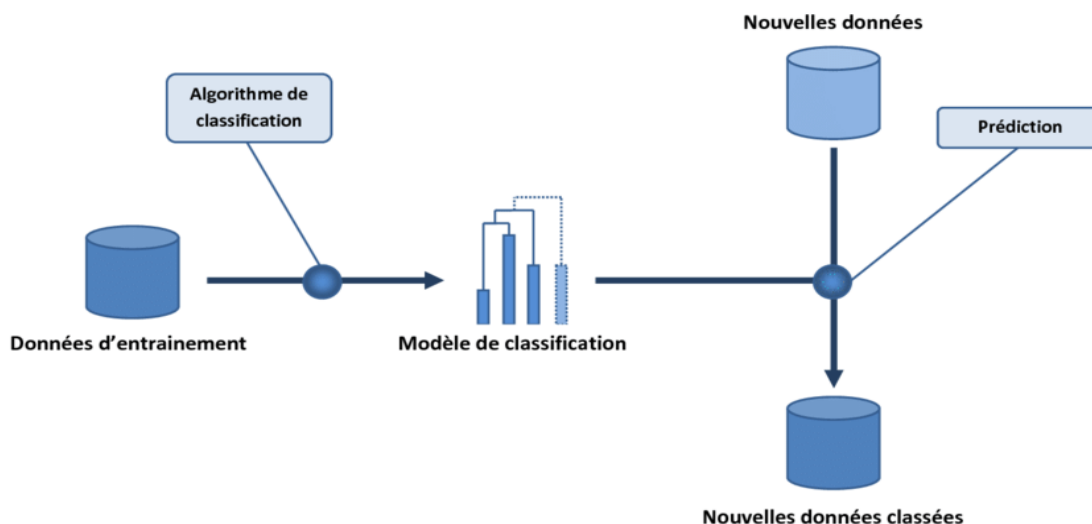


Figure II. 3 : Processus de classification supervisée et prédiction

Pour tester un processus apprentissage, comparer les algorithmes d'apprentissage et améliorer les performances sur un jeu de données, ils existent différentes techniques de test :

Holdout :

Le holdout, c'est quand vous divisez votre ensemble de données en un ensemble « train » et « test ». L'ensemble d'apprentissage est ceux sur quoi le modèle est formé, et l'ensemble de tests est utilisé pour voir dans quelle mesure ce modèle fonctionne sur les données. Une division commune lors de l'utilisation de la méthode de holdout consiste à utiliser 80% des

Chapitre II : Généralité sur l'intelligence artificielle et l'apprentissage automatique

données pour l'entraînement et les 20% restants des données pour les tests. Le principe de cette méthode de validation consiste à partitionner les données en exactement deux sous-ensembles d'un rapport spécifié pour la formation et la validation. [18]

K-cross validation :

Appelée estimation rotative la base de donnée D est divisé aléatoirement en k sous ensemble mutuellement exclusifs $D_1, D_2, \dots, D_k / D = D_1 \cup D_2 \dots \cup D_k$.

Le classificateur est testé k fois.

A chaque $t \in \{1, 2, \dots, k\}$, l'apprentissage se fait sur le data set $D-D_t$ et testé sur D_t . La performance finale P est la moyenne de toutes les performances P_t . Elle est calculée comme suit : [17]

$$P = \frac{\sum_{t=1}^k P_t}{k}$$

$$P_{t=1} = \frac{1}{N} \sum_{i=1}^n v(y - \hat{y})^2$$

La méthode leave-one-out : est tout simplement la méthode k-cross validation avec $k = n$, n étant le nombre d'instances. [17]

I.4.2 Les méthodes de Classification supervisée

I.4.2.1 Les réseaux de neurones artificiels

Le réseau de neurones artificiel (RNA) contient un ensemble de neurones artificiels fortement interconnectés inspirés des neurones biologiques du cerveau. Le but est d'imiter certaines fonctions du cerveau humain, tel que la mémorisation par association, l'apprentissage par exemple, etc.

Pour comprendre le fonctionnement des réseaux de neurones artificiels, il est utile de savoir comment fonctionnent les neurones naturels. Les neurones naturels reçoivent des signaux par des synapses situées sur les dendrites ou membrane du neurone. Quand les signaux reçus sont assez forts (surpassant un certain seuil), le neurone est activé et émet un signal à travers l'axone. Ce signal pourrait être envoyé à une autre synapse, et pourrait activer d'autres neurones. [19]

Chapitre II : Généralité sur l'intelligence artificielle et l'apprentissage automatique

Inspirés de ce mécanisme, les réseaux de neurones artificiels sont représentés par des nœuds qui constituent les <<neurones artificiels>>. Chaque neurone reçoit des signaux d'entrée (comme les synapses), pondérées par des poids (intensité des signaux respectifs) et puis calcule, par une fonction mathématique, l'activation (ou non) du neurone. Une autre fonction (qui peut être l'identité) calcule la valeur de sortie du neurone artificiel (parfois dépendant d'un certain seuil). Les réseaux de neurones artificiels combinent des neurones artificiels pour traiter l'information [19].

Le modèle réseau de neurones artificiels est entraîné avec des données d'exemple. Après l'entraînement, certains groupes de neurones seront activés en reconnaissant certains des patrons appris avec les données d'exemple, leur permettant ainsi de faire des prévisions de nouvelles données, encore non vues, en imitant ainsi le fonctionnement des neurones biologiques. Une des caractéristiques principales des réseaux de neurones est leur capacité d'apprendre des relations complexes non linéaires d'entrée et de sortie, en utilisant des procédures séquentielles. Ils sont, en outre, capables de s'adapter aux données.

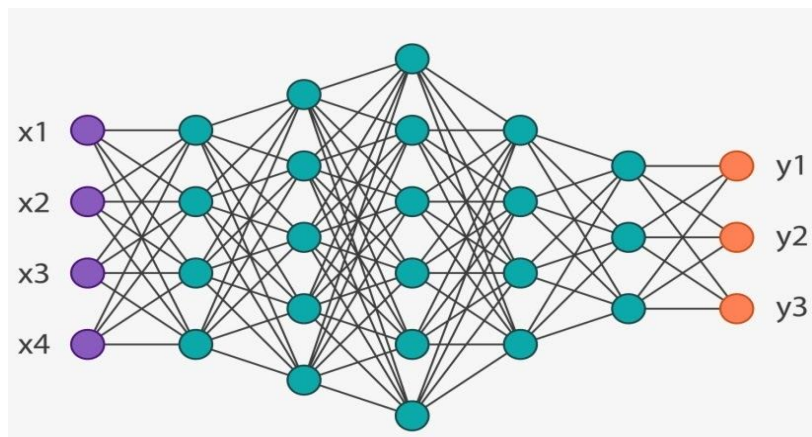


Figure II.3 : architecteur des réseaux de neurone

I.4.2.2 La machine à support vectoriel (SVM)

Les SVM utilisent une approche géométrique pour classer les données en deux catégories. En modélant les données comme des points (vecteurs) dans l'espace, elles construisent un hyperplan qui sépare les données dans chacune des catégories. Une fois la frontière de décision construite, la SVM sera capable de classer de nouvelles données en

Chapitre II : Généralité sur l'intelligence artificielle et l'apprentissage automatique

observant de quel côté de la frontière elles tombent, et en leur assignant la catégorie correspondante. [20]

La classification la plus optimale fonction peut être décidée en construisant un hyperplan qui maximise la marge entre deux ensembles de données et crée ainsi la plus grande distance possible entre les ensembles de données. Une visualisation de cette stratégie est illustrée à la figure II.4

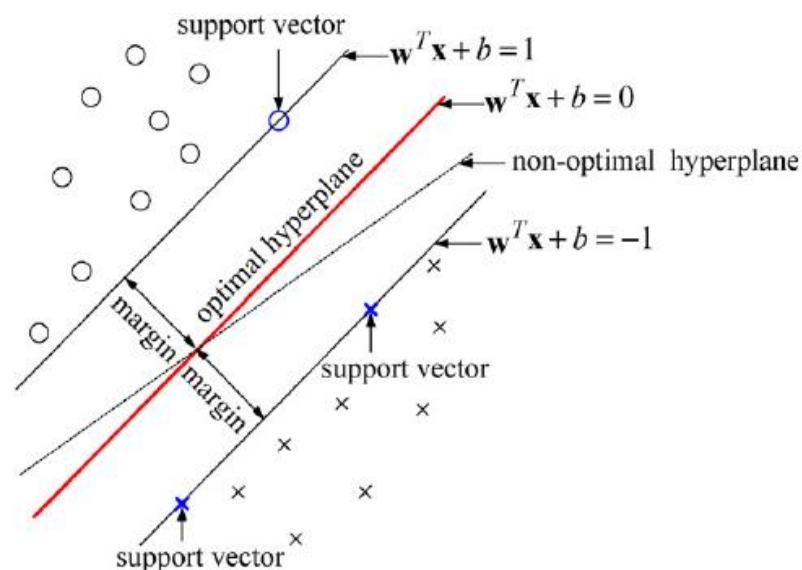


Figure II.4 : Création de l'hyperplan de séparation optimal à l'aide de vecteur de support

I.4.2.3. Arbre de décision

Les arbres de décision sont une des méthodes les plus connues en classification. Le principe des arbres de décision est de réaliser la classification d'un exemple par une suite de tests sur les attributs qui le décrivent. Concrètement, dans la représentation graphique d'un arbre.

1. Un nœud interne correspond à un test sur la valeur d'un attribut.
2. Une branche part d'un nœud et correspond à une ou plusieurs valeurs de ce test.

Chapitre II : Généralité sur l'intelligence artificielle et l'apprentissage automatique

3. Une feuille est un nœud d'où ne part aucune branche et correspond à une classe. Une règle de décision (de la forme si . . . alors . . .) est créée pour chaque chemin partant de la racine de l'arbre et Parcourant les tests (en faisant des conjonctions) jusqu'à la feuille qui est l'étiquette

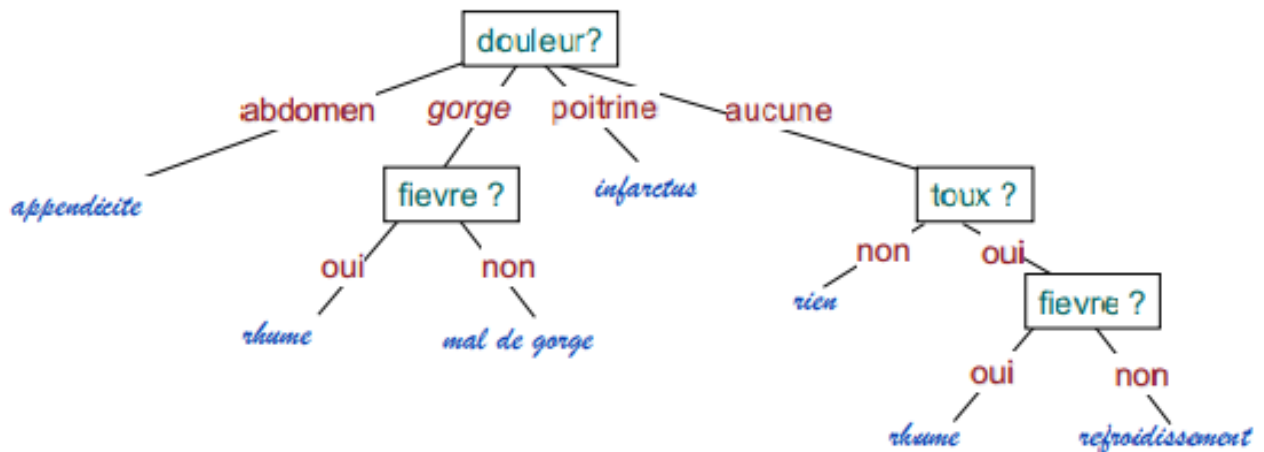


Figure II.5 : Un exemple de diagnostic d'un patient avec arbre de décision

I.4.2.4 Forêts aléatoires

Forêts aléatoires est une extension des arbres décision avec des particularités. Il s'agit d'une approche d'ensemble pour la classification et la régression qui fonctionne en construisant une multitude d'arbres de décision au moment de l'entraînement. La prédiction ou classification se fait alors en fonction d'un système de vote majoritaire au sein de ces différents arbres. Le principe de la forêt aléatoire est alors de chercher à tirer profit de cette instabilité en les agrégeant entre eux. [21]

La forêt aléatoire utilise de manière combinée les arbres de décision et la théorie statistique pour réduire la variance du classeur en calculant la moyenne d'un ensemble d'arbres de décision en générant des classeurs avec une très bonne capacité de prévision. [22]

I.4.2.5. Méthode des K plus proches voisins

L'algorithme des k-plus proches voisins (k-NN) se base sur les données en entier. En effet, pour une observation, qui ne fait pas partie des données, qu'on souhaite prédire,

Chapitre II : Généralité sur l'intelligence artificielle et l'apprentissage automatique

l'algorithme va chercher les k instances les plus proches de notre observation et choisir pour chaque observation la classe majoritaire parmi ses k plus proches voisins. La méthode k-NN est une technique d'apprentissage supervisé, et est considérée comme l'une des plus simples dans le domaine de la classification. Elle permet de classer une nouvelle observation (vecteur de caractéristiques extraites) en calculant la distance avec les données d'entraînement, et de prendre les k plus proches voisins (en termes de distance). Puis, observé la classe qui est majoritairement (voté) représentée parmi les k-plus proches voisins et D'assigne cette classe à la nouvelle observation

Le temps

d'apprentissage de l'algorithme k-NN soit court, l'algorithme k-NN calcule la distance entre les points de données. Pour cela, nous utilisons la formule de la distance euclidienne : [24]

$$d(p, q) = d(q, p) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2}$$
$$d(p, q) = d(q, p) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$$

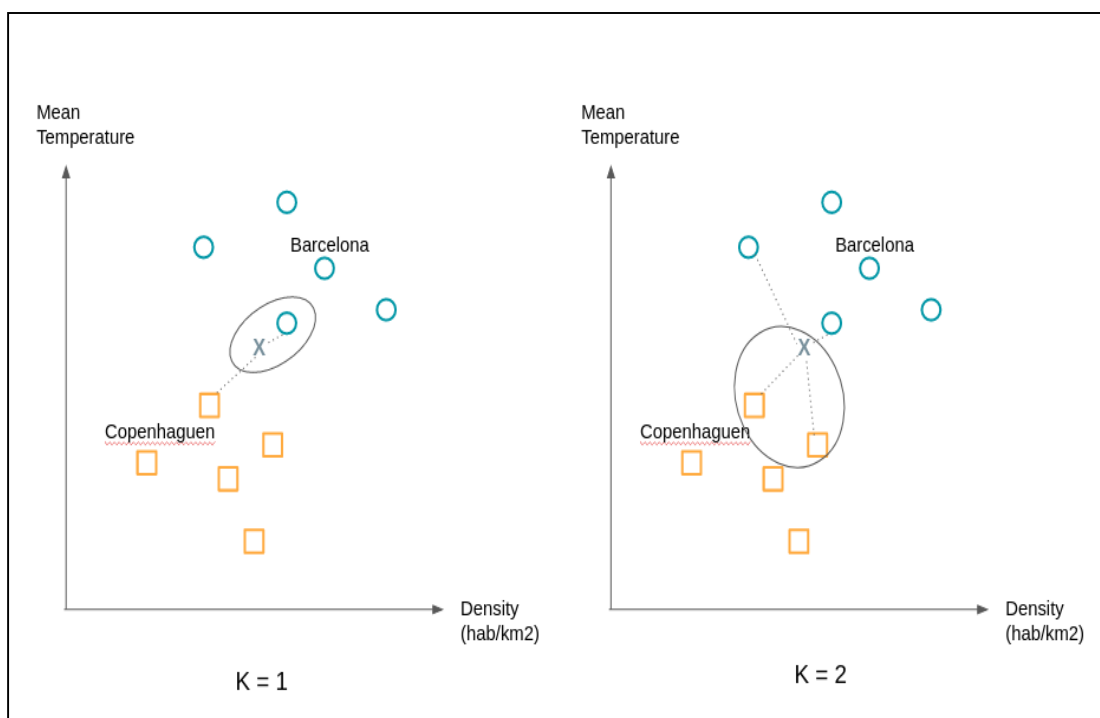


Figure II.6 : Visualisation d'un exemple de classification en utilisant

Chapitre II : Généralité sur l'intelligence artificielle et l'apprentissage automatique

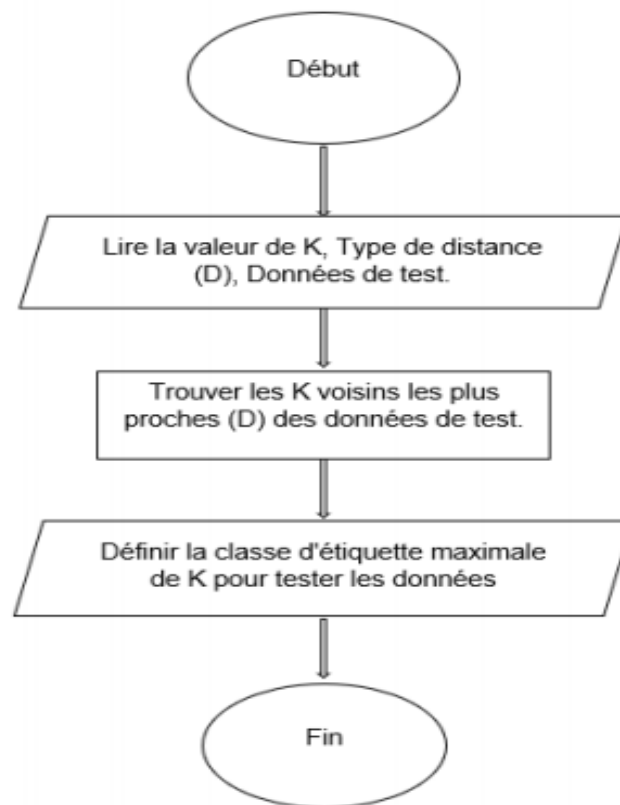


Figure II.7 : Organigramme de l'algorithme KNN

I.4.3. L'évaluation de performance des modèle d'apprentissage

I.4.3.1 Taux de classification (accuracy) et Taux d'erreur (error)

Si nous pouvons tester le classificateur sur tous les exemples possibles, nous consacrons exactement quel est son taux de classification. Malheureusement, il est généralement impossible d'effectuer de tels tests ; un estimé devra donc être utilisé à sa place. Pour un ensemble de Q_t exemples de test, l'estimation du taux de classification est déterminée par la réalisation des tests sur tous les exemples disponibles. Il s'agit ensuite de calculer soit le rapport des exemples bien classés (N_{juste}), soit de celui des exemples mal classés (N_{faut}). Erreur accuracy désigne le pourcentage d'exemple mal classée

Chapitre II : Généralité sur l'intelligence artificielle et l'apprentissage automatique

$$accuracy = \frac{NJUST}{QT}$$

$$Erreur\ accuracy = \frac{Nfaut}{QT} = 1 - accuracy$$

I.4.3.2. Matrice de confusion

Il s'agit d'un tableau de taille n x n pour visualiser les résultats des modèles prédictifs pour les problèmes de classification. Où 'n' est le nombre de classes dans nos ensembles de données. Dans cette matrice on croise les classes cibles réelles avec les classes prédites obtenues. Ceci nous donne le nombre ou le pourcentage d'instances qui sont correctement classées et le nombre ou pourcentage d'instances mal classées.

I.5. Conclusion

Dans ce chapitre, nous avons introduit les principes de l'intelligence artificielle et de l'apprentissage automatique, il est connu que l'apprentissage automatique nécessite des données volumineuses pour apprendre, mais dans certains cas, le grand volume de données est une obstruction, car il entraîne un temps de calcul long, ce qui affecte négativement les capacités de l'ordinateur. Il peut également faire échouer le système dans l'apprentissage, donc les performances optimales n'ont pas atteint, pour cela en a besoin de faire la minimisation des données et la suppression des données qui sont pas importants, cette technique est s'appelle la sélection des caractéristiques.

Nous mentionnerons dans le prochain chapitre les notions d'optimisation et de la sélection des caractéristiques.

Chapitre III : Notion d'optimisation et de la sélection des caractéristiques

II.1. Introduction

L'optimisation est un problème très ancien qui a donné lieu à de nombreuses méthodes de calcul. Un grand nombre de problèmes peuvent, en effet, être décrits sous la forme de problèmes d'optimisation. C'est le cas, par exemple, pour les problèmes de la sélection des caractéristiques. Pour éviter la « problème de la dimensionnalité » et pour réduire le coût de calcul, il est toujours préférable de se limiter à un petit nombre des caractéristiques discriminants. Le problème qui se pose alors est de choisir le sous-ensemble d'attributs permettant d'obtenir les meilleures performances et la meilleure classification.

Dans ce chapitre nous présentons quelques notions d'optimisation et comment voir le problème de la sélection comme un problème d'optimisation, et, nous présenterons le processus général de sélection des caractéristique en identifiant ses différentes étapes de la sélection.

Enfin nous avons présenté l'algorithme d'optimisation (aco) pour la sélection des caractéristiques.

III.2. Quelque notion sur l'optimisation

III.2.1. Définition d'Optimisation

L'optimisation est un aspect fondamental de l'ingénierie et de la résolution de problèmes. L'objectif de l'optimisation est de chercher les valeurs d'un ensemble de paramètres pour maximiser ou minimiser les fonctions objectives soumises à certaines contraintes. Un choix des valeurs, pour l'ensemble de paramètres, qui satisfont toutes les contraintes, est appelé une solution faisable. Les solutions faisables avec des valeurs de la fonction objective qui sont meilleures que les valeurs de toutes les autres solutions possibles, sont appelées les solutions optimales. [24].

III.2.2. Formule de l'optimisation

Un problème d'optimisation en général est défini par un espace de recherche S et une fonction objectif f . Le but est de trouver la solution $s^* \in S$ de meilleure qualité $f(s^*)$. Suivant le problème posé, on cherche soit le minimum soit le maximum de la fonction f . L'équation (1)

Chapitre III : Notion d'optimisation et de la sélection des caractéristiques

montre bien l'aspect de minimisation, maximiser une fonction f étant équivalent à minimiser f . [25]

$$s = \min(f(s) \mid s \in S)$$

De plus, un problème d'optimisation peut présenter des contraintes d'égalité et/ou d'inégalité sur les solutions candidates $s \in S$. [25]

- Quand il s'agit d'une seule fonction, nous parlons d'optimisation mono-objectif. Dans le cas contraire, nous parlons d'optimisation multi-objectif.
- Quand nous essayons de minimiser une fonction, nous appelons la valeur de la fonction : la fonction de coût.
- Quand nous essayons de maximiser une fonction, nous appelons la valeur de la fonction : le fitness.

III.3. Les domaines d'application de l'optimisation

L'optimisation est essentiellement un outil d'aide à la décision au sein de l'entreprise, mais aussi pour des individus. Le terme optimal est souvent trompeur. Ce n'est pas un jugement de valeur absolu. C'est plutôt une information sur l'approche méthodologique utilisée.

Voici quelques domaines dans lesquels la prise de décision fait couramment appel à la résolution de problèmes d'optimisation :

- Conception de nouveaux systèmes dimensionnement, localisation
- Organisation d'activité Gestion de ressources Gammes et procédures Rentabilisation des investissements
- Commande de systèmes Stabilisation, Suivi de trajectoires
- Surveillance, supervision, Détections de dysfonctionnements, diagnostics, réparation, maintenance, remplacement préventif.
- en recherche opérationnelle (problème de transport, économie, gestion de stocks...).

III.4. Notion sur la sélection des caractéristiques

III.4.1 Définition de la sélection des caractéristiques

La sélection d'attributs, de caractéristiques ou, est un problème difficile qui a été étudié depuis les années 70. La sélection de caractéristiques est un processus de recherche ou une technique utilisée pour choisir les caractéristiques, les variables ou les mesures les plus intéressantes, pertinentes ou informatives d'un système donné, dans le but de réaliser la tâche pour laquelle il a été conçu. Dans le domaine d'apprentissage automatique ou, plus précisément, celui de la classification, certaines caractéristiques non pertinentes et/ou redondantes, existant généralement dans les données d'apprentissage, non seulement rendent l'apprentissage plus difficile, mais dégradent aussi les performances de généralisation des modèles d'apprentissage. [26]

La définition proposée par Pudil et al. Dans [27] est la suivante : étant donnée une fonction permettant de mesurer la qualité d'un sous-ensemble de caractéristiques, la sélection des caractéristiques est réduite au problème de la recherche du sous-ensemble optimal par rapport à cette mesure.

Dans [28] et [29], l'énoncé de la sélection de caractéristiques proposée est la suivante : étant donné un ensemble de dimension n , il faut sélectionner le sous-ensemble de dimension m tel que $m < n$, conduisant au taux d'erreur le plus faible.

III.4.2 Objectifs de la sélection des caractéristiques

Dans un problème de reconnaissance de forme, la performance du système de reconnaissance devient mauvaise et le temps de calcul augmente à mesure que le nombre de caractéristiques augmente. Dans certains cas, il existe quelques caractéristiques qui n'aideront pas à discriminer entre les classes. Autrement dit, il existe des caractéristiques redondantes ou non pertinentes inutiles dans la classification de l'objet d'où l'utilité d'effectuer la sélection des caractéristiques les plus pertinentes afin de rendre l'apprentissage simple et rapide. [30]

En effet, les principaux objectifs de la sélection des caractéristiques sont :

Identifier les caractéristiques pertinentes

Réduire les bases d'apprentissage et de test

Chapitre III : Notion d'optimisation et de la sélection des caractéristiques

Réduire la dimension de l'espace d'entrée en éliminant les informations non pertinentes et redondante

Améliorer les performances en classification

Réduire le temps d'apprentissage

Améliorer la vitesse de la classification.

III.4.3 La sélection vue comme un problème d'optimisation

Le problème de la sélection d'un sous-ensemble d'attributs peut être vu comme une recherche dans un espace d'hypothèses (appelé ensemble de solutions possibles) [41]. Étant donné un ensemble initial X de n attributs, la sélection d'un "bon" sous-ensemble d'attributs nécessite d'examiner potentiellement $2^n - 1$ sous-ensembles possibles. La qualité d'un sous-ensemble sélectionné est évaluée selon un critère de performance que l'on notera J . Dans le cas d'un problème de classification supervisée, ce critère est très souvent la précision d'un classifieur construit à partir de l'ensemble d'attributs sélectionnés.

La recherche d'un sous-ensemble d'attributs, optimal pour le critère J que l'on s'est donné, est alors un problème NP-difficile [Davies et Russell, 1994; Cotta et Moscato, 2003]. Plusieurs approches peuvent être envisagées pour contourner cette difficulté. Elles sont formalisées dans la définition suivante :

[Molina et al., 2002a] Soit X un ensemble d'attributs. Soit J une mesure d'évaluation qui attribue à tout sous-ensemble de X un score : $J: \bar{X} \subseteq X \rightarrow \mathbb{R}$:

J doit être optimisée (maximisée ou minimisée suivant la nature de J), on supposera dans la suite que J doit être maximisée .

La sélection d'un sous-ensemble d'attributs peut se faire suivant un des schémas suivants :

- Nombre d'attributs fixé: Pour un nombre m fixé, avec $m < n$, on cherche à trouver $\bar{X} \subset X$ tel que $|\bar{X}| = m$ et que $J(\bar{X})$. Soit maximum.
- Seuil de performance fixé: On se donne une valeur seuil J_{opt} , c'est-à-dire, le minimum acceptable pour J , et on cherche à trouver $\bar{X} \subset X$ tel que le cardinal de \bar{X} soit le plus petit possible et que $J(\bar{X}) \geq J_{opt}$

Chapitre III : Notion d'optimisation et de la sélection des caractéristiques

– Compromis performance et nombre d'attributs. Trouver un compromis entre le fait de minimiser le nombre d'attributs $|\bar{X}|$ et le fait d'optimiser $J(\bar{X})$.

III.4.4. Procédure de la sélection des caractéristiques

La procédure de sélection de variables tente de sélectionner le sous ensemble le plus petit selon deux principaux critères :

La précision de la classification ne se dégrade pas,

La distribution des classes est proche de la distribution originale.

Trois questions essentielles devraient se poser à ce stade afin de définir les éléments essentiels d'une procédure de sélection de variables : [31]

Comment mesurer la pertinence des variables ?

Comment former le sous-ensemble optimal ?

Quel est le critère d'optimalité à utiliser ?

La pertinence d'une variable se définit selon deux critères. Soit son pouvoir discriminant permettant une meilleure séparation entre les différentes classes dans le cadre d'une classification. Soit son pouvoir prédictif garantissant une meilleure qualité de prédiction dans le cadre d'une régression. Il serait alors naturellement nécessaire de définir une mesure de pertinence ou un critère d'évaluation permettant de quantifier l'utilité d'une variable ou d'un ensemble de variables.

Une procédure de sélection d'attributs peut être décortiquée en quatre étapes typiques illustrées dans la figure. [32]

-Procédure de génération,

-Fonction d'évaluation,

-Critère d'arrêt,

-Procédure de validation.

Chapitre III : Notion d'optimisation et de la sélection des caractéristiques

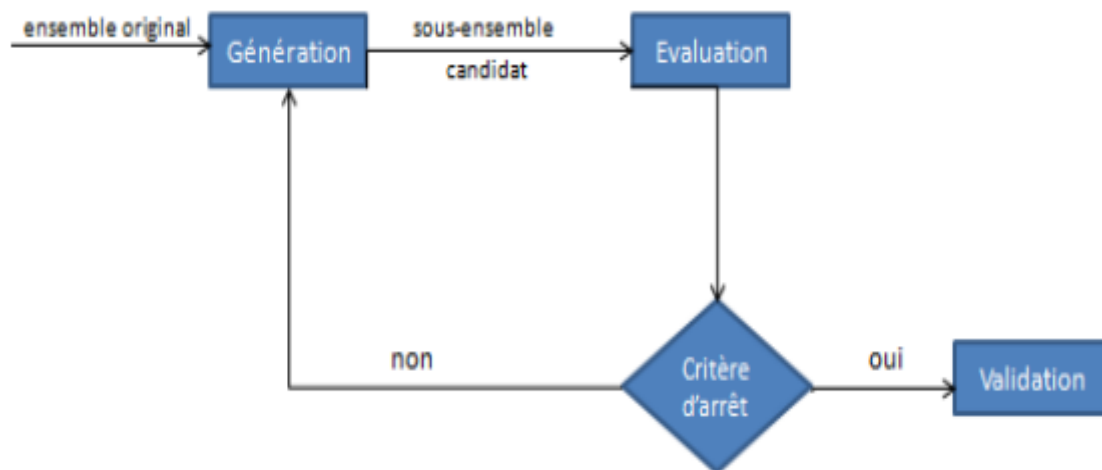


Figure III.1: Procédure de sélection des caractéristiques

III.4.4.1. Les procédures de recherches

En général, les stratégies de recherche pour la génération des sous-ensembles dans le processus de sélection des attributs peuvent être classées en trois catégories : exhaustive, heuristique et aléatoire [33].

III.4.4.1.1. Génération exhaustive

Dans cette approche, une recherche exhaustive sur tous les sous-ensembles de caractéristiques est effectuée afin de sélectionner le "meilleur" sous-ensemble de caractéristiques. Cette stratégie de recherche garantit de trouver le sous-ensemble optimal. Le problème majeur de cette approche est que le nombre de combinaisons croît exponentiellement en fonction du nombre de caractéristiques. Pour un ensemble de N caractéristiques, et quand N devient grand, les 2^N combinaisons possibles rendent la recherche exhaustive impossible (problème NP-complet, Blum et Rivest [34]).

III.4.4.1.2. Génération heuristique

Elle regroupe les algorithmes itératifs pour lesquels chaque itération permet de sélectionner ou rejeter une ou plusieurs caractéristiques. Ils ont l'avantage d'être assez simples et rapides. En revanche, ils ne permettent de parcourir généralement qu'un petit sous-espace de l'espace total des possibilités. Dans la littérature, les trois sous-catégories les plus connues de cette approche

Chapitre III : Notion d'optimisation et de la sélection des caractéristiques

sont :

Forward (approche ascendante) : cette approche part d'un ensemble de caractéristiques vide auquel, à chaque itération sont ajoutées une ou plusieurs caractéristiques.

Backward (approche descendante) : cette approche procède à l'inverse de "Forward". L'ensemble de départ représente l'ensemble total des caractéristiques puis à chaque itération, une ou plusieurs caractéristiques sont supprimées.

Stepwise (approche hybride) : cette dernière approche consiste à combiner les deux précédentes par l'ajout ou la suppression successive des caractéristiques de l'ensemble des caractéristiques courante à chaque itération.

III.4.4.1.3. Génération aléatoire

Pour un ensemble de données et une initialisation particulière, une stratégie de recherche heuristique retourne toujours le même sous-ensemble, ce qui la rend très sensible au changement de l'ensemble de données. La procédure de recherche aléatoire (appelée aussi stochastique ou non-déterministe) consiste à générer aléatoirement un nombre fini de sous-ensembles de caractéristiques afin de sélectionner le meilleur. En outre, les stratégies de recherche aléatoires convergent en général rapidement vers une solution "semi-optimale", ce qui est préférable pour éviter le phénomène de sur-apprentissage. [35]

III.4.4.2. La fonction d'évaluation

Le but d'une fonction d'évaluation est de mesurer la capacité d'une variable ou d'un ensemble de variables exogènes à distinguer les classes de la variable endogène. L'optimalité d'un sous-ensemble est relative à la fonction d'évaluation utilisée. Dash et Liu considèrent que ces fonctions ou critères peuvent être regroupées en cinq catégories qui sont les suivantes :

- **Critère d'information :** C'est la quantité d'information apportée par une variable sur la variable endogène. La variable, ayant le gain d'information le plus élevé, sera préférée aux autres variables. Le gain d'information est la différence entre l'incertitude a priori et l'incertitude a posteriori.

- **Critère de distance :** Ces mesures s'intéressent au pouvoir discriminant d'une variable

- **Critère d'indépendance :** Ils regroupent toutes les mesures de corrélation ou d'association. Ils permettent de calculer le degré avec lequel une variable exogène est associée à une variable endogène.

Chapitre III : Notion d'optimisation et de la sélection des caractéristiques

- **Critère de consistance** : Ils sont liés au biais des variables minimums. Ces mesures recherchent l'ensemble de variables le plus petit qui satisfait un pourcentage d'inconsistance minimum défini par l'utilisateur. Deux objets sont inconsistants si leurs modalités sont identiques et s'ils appartiennent à deux classes différentes. Ces mesures peuvent permettre de détecter les variables redondantes.
- **Critère de précision** : Ils utilisent le classifieur comme fonction d'évaluation. Le classifieur choisit, parmi tous les sous-ensembles de variables, celui qui est à l'origine de la meilleure précision prédictive. (Meilleur accuracy).

Critère d'arrêt :

Le critère d'arrêt définit les bornes de la recherche.

Critère d'arrêt peut être un nombre fixé d'attributs à retenir, un temps de calcul fixé, un nombre d'itérations fixé, lorsqu'il n'y a plus d'amélioration de précision, (pas de possibilité de trouver un sous-ensemble meilleur que le sous-ensemble actuel).

III.4.4.3. Validation du sous ensemble

La sélection des caractéristiques est considérée comme une phase de prétraitement pour faire la classification, Les sous-ensembles sélectionnés sont validés grâce aux performances d'apprentissage utilisé par la suite, et cela par la comparaison entre les taux de réussite avant et après la sélection.

III.4.5. Les méthodes de sélection d'attributs

Les techniques de sélection d'attributs sont divisibles en trois catégories selon leur critère d'évaluation et la manière dont elles interagissent avec le classifieur.

III.4.5.1 Les méthodes filtres

Cette méthode évalue la pertinence d'une caractéristique selon des mesures qui reposent sur les propriétés des données d'apprentissage avant la phase d'apprentissage. En d'autres termes, l'évaluation se fait généralement indépendamment d'un classificateur [36]. Les méthodes qui se basent sur ce modèle pour l'évaluation des caractéristiques, utilisent souvent une approche heuristique comme stratégie de recherche. [34] La procédure du modèle "filtre" est illustrée par la figure

Chapitre III : Notion d'optimisation et de la sélection des caractéristiques

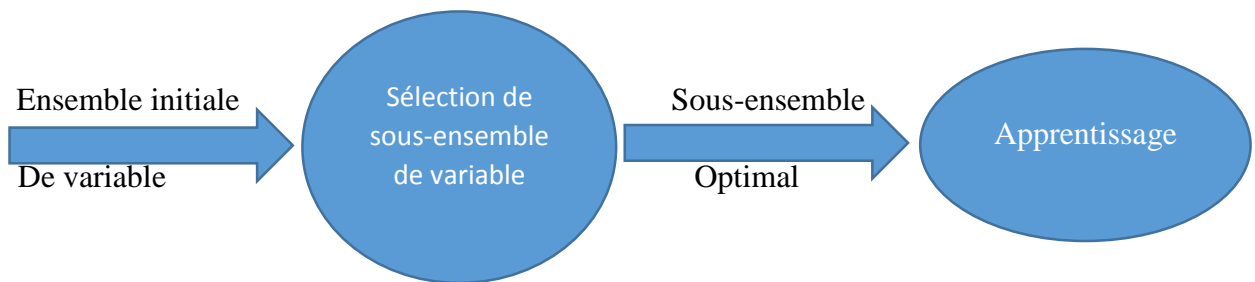


Figure III.2 : La procédure du méthode "filtre"

III.4.5.2 La méthode Embedded (embarquée)

Ces méthodes effectuent la sélection en parallèle avec la procédure de classification ou de la régression. Le sous-ensemble sélectionné sera choisi de façon à optimiser le critère d'apprentissage utilisé. [38] L'avantage principal de cette technique est qu'elle est plus rapide que la technique "Wrapper" :

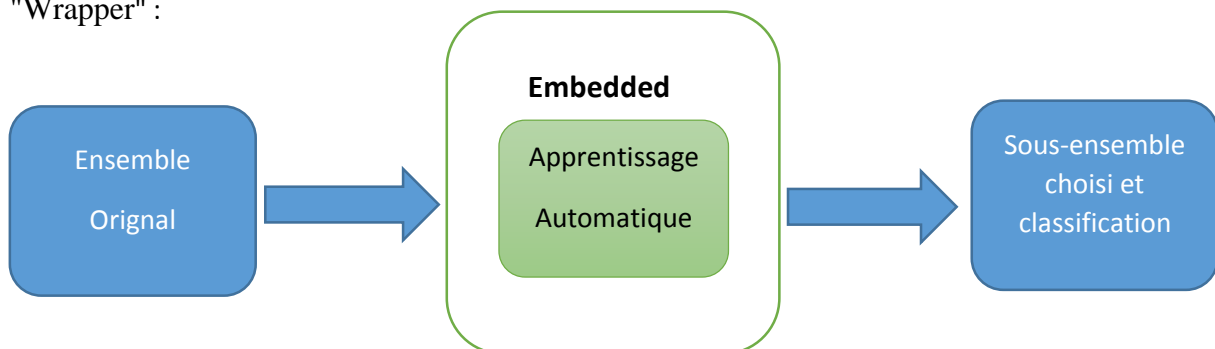


Figure III.3 : La procédure de la méthode Embedded (embarquée)

III.4.5.3. Wrapper méthode

Le principal inconvénient de la technique "filter" provient du fait qu'elle ignore l'influence des attributs sélectionnés sur la performance du classificateur. Dans [37], Kohavi et John proposent la technique wrapper pour résoudre ce problème. Le sous ensemble sélectionné sur cette technique est dépend de l'algorithme de classification (évalue un sous -ensemble de caractéristique en utilisant l'algorithme de classification. Cette technique produit une précision plus élevée, La procédure du modèle "wrapper" est illustré par la Figure

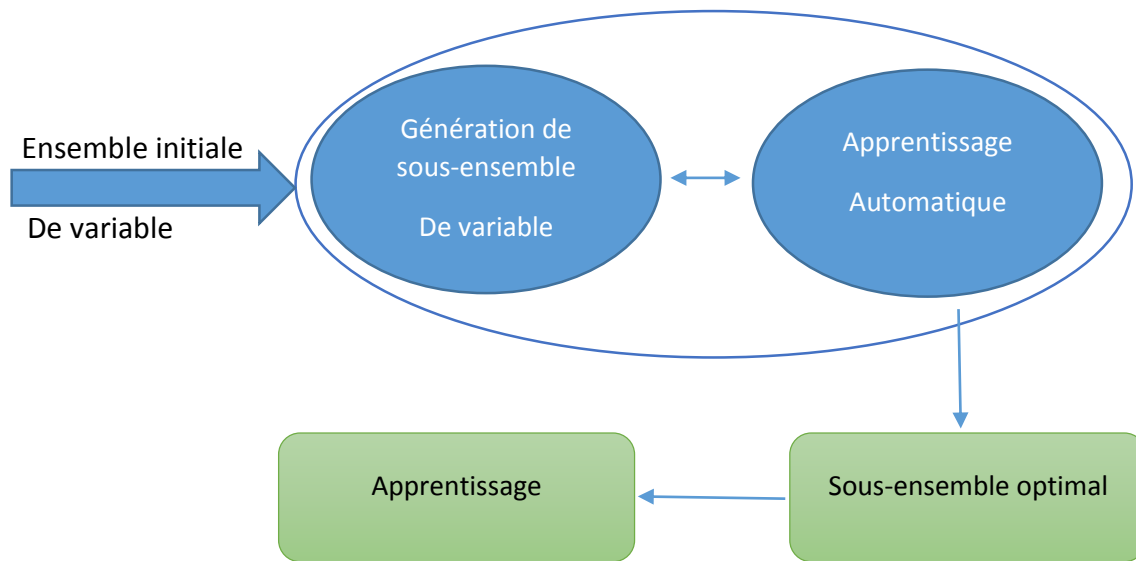


Figure III.4 : La procédure de la méthode wrapper

III.5. L'algorithme d'optimisation (aco) pour la sélection des Caractéristiques

L'algorithme colonie des fourmis (aco) a été inspiré par les études sur le comportement des fourmis qui sont capables de résoudre collectivement des problèmes complexes, comme trouver le plus court chemin pour aller chercher la nourriture. Cette algorithmique a été initialement proposée dans [39,40], comme une approche pour résoudre des problèmes d'optimisation. L'idée est de représenter le problème à résoudre sous la forme de la recherche d'un meilleur chemin dans un graphe, puis d'utiliser des fourmis artificielles pour rechercher de bons chemins dans ce graphe. Le comportement des fourmis artificielles est inspiré des fourmis réelles.

III.5.1 Le comportement des fourmis pour chercher la nourriture

Les fourmis marchant du nid à la source de nourriture (ce qui, dans un premier temps, se fait essentiellement d'une façon aléatoire) et déposent au passage sur le sol une substance odorante appelée phéromone pour créer une piste chimique. Quand les fourmis choisissent leur chemin, elles ont tendance à choisir la piste qui porte la plus forte concentration de phéromone. Les fourmis les plus rapidement arrivées au nid, après avoir visité la source de nourriture, sont celles qui empruntent les branches les plus courtes. Ainsi, la quantité de phéromone présente sur le plus court trajet est légèrement plus importante que celle présente sur le chemin le plus

Chapitre III : Notion d'optimisation et de la sélection des caractéristiques

long. Or, une piste présentant une plus grande concentration en phéromone est plus attirante pour les fourmis, elle a une probabilité plus grande d'être empruntée. La piste courte va alors être plus renforcée que la longue, et sera choisie par la grande majorité des fourmis.

La phéromone est caractérisée par l'évaporation au fil du temps. Lorsque les fourmis abandonnent un chemin, la phéromone s'évapore, et ainsi les fourmis ne sont pas attirées par le mauvais chemin et évitent la désinformation.

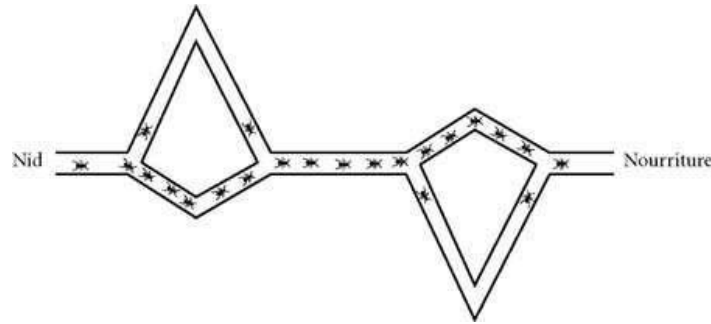


Figure III.5 : schéma de comportement des fourmis pour choisir le plus court chemin de recherche de nourriture

III.5.2 Le principe de l'algorithme des fourmis (ant colony optimisation)

Le premier algorithme qui s'inspire de cette analogie a été proposé en 1996 par Colomni, Dorigo et Maniezzo

À l'origine, l'optimisation par colonie de fourmis a été proposée en 1996 par Colomni, Dorigo et Maniezzo pour résoudre le problème du voyageur de commerce (VC). L'objectif de ce problème est de trouver la tournée la plus courte pour un ensemble de villes données.

Les fourmis artificielles cherchent d'un meilleur chemin (le plus court) dans un graphe pour visiter tous les villes.

Chaque fourmi k parcourt le graphe et construit un trajet de longueur $n = |N|$. Pour chaque fourmi K , le trajet d'une ville i à une ville j dépend de :

J_i^K : La liste des villes déjà visitées

d_{ij} : distance entre deux ville

Chapitre III : Notion d'optimisation et de la sélection des caractéristiques

$\eta_{ij} = \frac{1}{d_{ij}}$: visibilité (Cette information est utilisée pour diriger les fourmis vers des villes proches et ainsi, éviter de trop longs déplacements)

τ_{ij} : La quantité de phéromone déposée sur l'arête reliant deux villes. Cette quantité définit l'attractivité d'une piste (intensité de la piste)

La règle de déplacement est la suivante :

$$p_{ij}^k(t) = \begin{cases} \frac{(\tau_{ij}(t))^\alpha (\eta_{ij})^\beta}{\sum_{l \in J_i^k} (\tau_{il}(t))^\alpha (\eta_{il})^\beta} & \text{si } j \in J_i^k \\ 0 & \text{si } j \notin J_i^k \end{cases}$$

Où α et β sont deux paramètres contrôlant l'importance relative de l'intensité et de la visibilité. Après un tour complet, chaque fourmi dépose une quantité de phéromone $\Delta\tau_{ij}^k(t)$ sur l'ensemble de son parcours. Cette quantité dépend de la qualité de la solution trouvée, et elle est définie par

$$\Delta\tau_{ij}^k(t) = \begin{cases} \frac{Q}{L^k(t)} & \text{si } (i,j) \in T^k(t) \\ 0 & \text{si } (i,j) \notin T^k(t) \end{cases}$$

$T^k(t)$: le trajet effectué par la fourmi k à l'itération t ,

$L^k(t)$: la longueur de $T^k(t)$ et Q est un paramètre fixé.

Enfin, il est nécessaire d'introduire un processus d'évaporation des phéromones. En effet, pour éviter de se faire piéger dans des solutions sous-optimales, il est nécessaire qu'une fourmi "oublie" les mauvaises solutions. La règle de mise à jour est donc.

$$T_{ij}(t+1) = (1-p) \cdot T_{ij}(t) + \Delta t_{ij}(t)$$

où $\Delta t_{ij}(l) = \sum_{k=1}^m \Delta t_{ij}^k(t)$ et m est le nombre de fourmis.

p : coefficient d'évaporation des phéromones

Chapitre III : Notion d'optimisation et de la sélection des caractéristiques

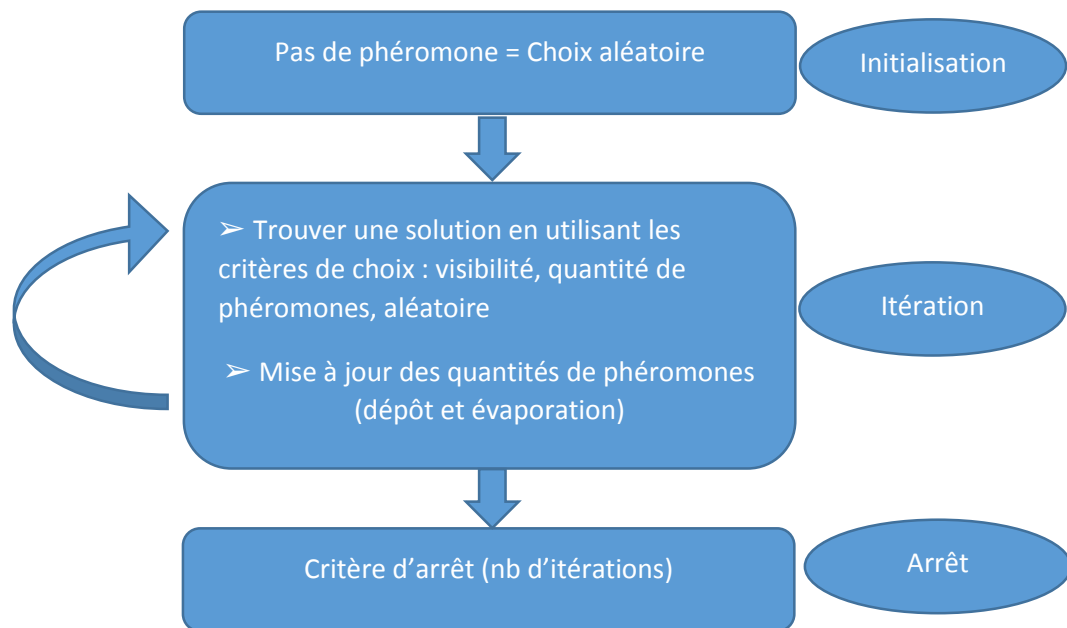


Figure III.6 : principe de l'algorithme des fourmis

Le problème de sélection d'attributs peut être considéré comme un problème d'optimisation, l'algorithme des colonies de fourmis ont été largement appliquées pour la sélection des caractéristiques avec même principe de problème de VS avec quelque point différent.

Dans TSP, une fourmi doit passer par toutes les villes, dans le problème de sélection d'attributs, une fois que l'ensemble des sommets C parcourus par la fourmi satisfait l'objectif de la sélection, les fourmis arrêtent le déplacement, puis l'ensemble est une solution.

La tâche de sélection de caractéristiques peut être reformulée en un problème ACO convenable. ACO exige qu'un problème soit représenté comme un graphe. Ici, les nœuds représentent les caractéristiques, les arêtes entre eux indiquent le choix de la prochaine caractéristique. La recherche du sous-ensemble optimal de caractéristiques est alors un parcours des fourmis dans le graphe avec un nombre minimal de nœuds visités satisfaisant le critère d'arrêt du parcours (par exemple, une précision de classification suffisamment élevée a été atteinte avec ce sous-ensemble).

Dénotant la probabilité d'une fourmi sur la caractéristique i choisissant à voyager à la caractéristique j à l'instant t par :

Chapitre III : Notion d'optimisation et de la sélection des caractéristiques

$$p_{ij}^k(t) = \begin{cases} \frac{(\tau_{ij}(t))^\alpha (\eta_{ij})^\beta}{\sum_{l \in J_i^k} (\tau_{il}(t))^\alpha (\eta_{il})^\beta} & \text{si } j \in J_i^k \\ 0 & \text{sinon} \end{cases}$$

Où k est le nombre de fourmis, J_i^k l'ensemble des caractéristiques non visitées par la fourmi k , η_{ij} est l'opportunité de choisir la caractéristique j lorsque la fourmi est sur la caractéristique i et $\tau_{ij}(t)$ est la quantité de phéromone virtuelle sur l'arête (i, j) . Les valeurs de α et β sont choisies expérimentalement.

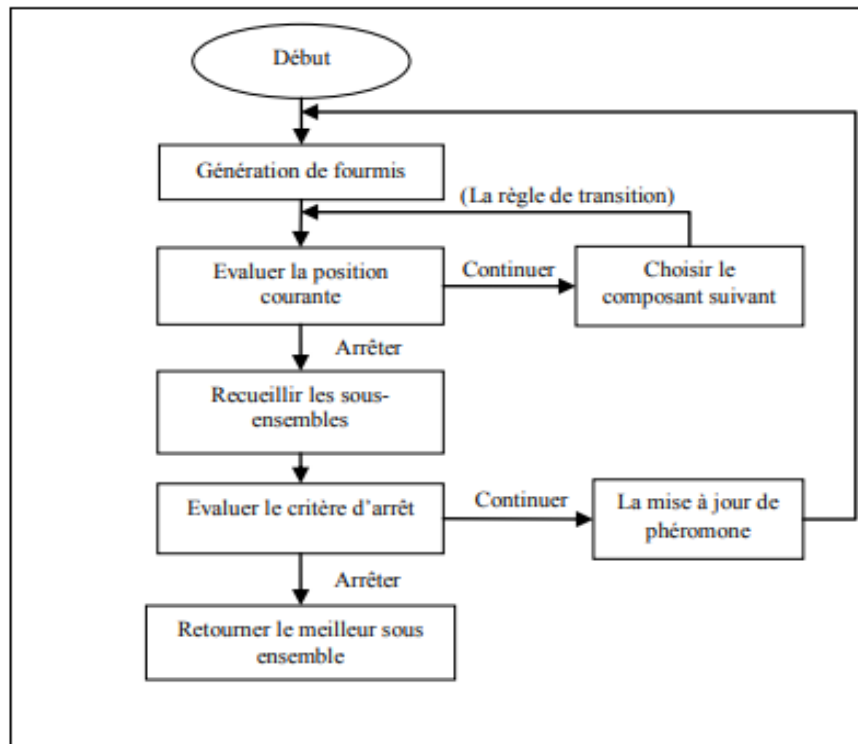


Figure III.7: Le processus global de sélection d'attributs par algorithme des fourmis (ACO)

III.5. Conclusion

Ce chapitre décrit les principes de l'optimisation et de la sélection, et la relation entre les deux. Nous avons également mentionné dans le chapitre les concepts et les méthodes de la sélection des caractéristiques, ensuite nous avons présenté l'algorithme des fourmis (ant colony) utilisé pour l'optimisation. Vers sa fin, Nous avons présenté une brève description de la manière avec laquelle l'algorithme des fourmis sont utilisés pour la sélection d'attributs.

Chapitre IV : application d'algorithme de fourmis et KNN pour la sélection et la classification des caractéristiques

IV.1. Introduction

En raison de l'importance des systèmes hydrauliques dans l'industrie, nous devons choisir une technique très efficace pour le diagnostic des systèmes hydrauliques.

Pour atteindre cet objectif, nous avons utilisé la méthode d'intelligence artificielle KNN sur un système hydraulique complexe qui contient de grand nombre de capteurs.

Nous avons amélioré les performances de notre technologie par l'utilisant d'algorithme des fourmis pour Résoudre le problème de l'optimisation de nombre des captures et pour la sélection des caractéristiques pour obtenir de meilleures performances.

Dans ce chapitre, nous présentons la méthodologie d'application des deux algorithmes pour diagnostiquer notre système avec interprétation et comparaison des résultats obtenus pour démontrer l'efficacité de la technique utilisée par rapport à l'autre méthode.

IV.2. Description de système et de base des donné utilisé

L'ensemble de données a été obtenu expérimentalement avec un banc d'essai sur un système hydraulique se compose d'un travail primaire (Fig. 2a) et un circuit secondaire de refroidissement-filtration (Fig. 2b) qui sont connecté via le réservoir d'huile [1].

Quatre défauts sont imposés au système hydraulique, Le système répète cycliquement le constant cycle de charge (durée 60 secondes) et mesure les valeurs des pressions, température, le débit volumique, l'énergie et les vibrations a l'laide de 17 capture

Tandis que l'état de quatre composants hydrauliques (refroidisseur, valve, pompe et accumulateur) est varié quantitativement.

Chapitre IV : application d'algorithmes de fourmis et KNN pour la sélection et la classification des caractéristiques

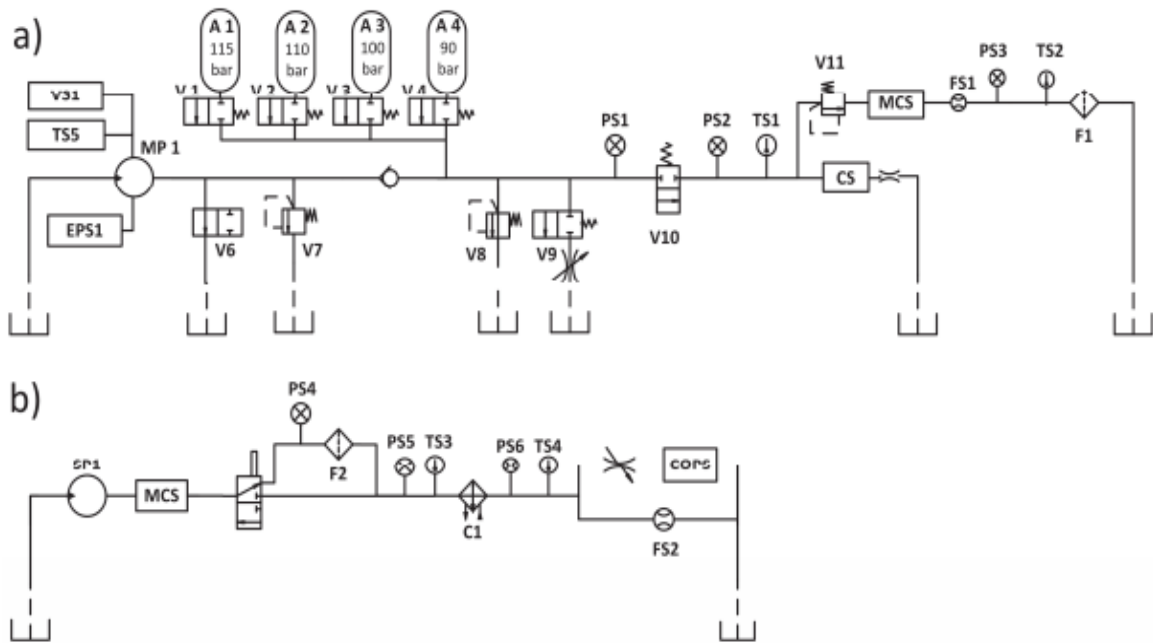


Figure IV.1 : schéma de notre système hydraulique

L'ensemble de données contient des données brutes de capteur de processus (c'est-à-dire sans extraction de caractéristiques) qui sont structurées sous forme de matrices (délimitées par des tabulations) avec les lignes représentant les cycles et les colonnes les points de données dans un cycle. Les capteurs concernés sont [2] :

Capteur Grandeur physique Unité Taux d'échantillonnage

PS1 Pression bar 100 Hz

FS2 Débit l/min 10 Hz

PS2 Pression bar 100 Hz

TS1 Température °C 1 Hz

PS3 Pression bar 100 Hz

TS2 Température °C 1 Hz

PS4 Pression bar 100 Hz

TS3 Température °C 1 Hz

PS5 Pression bar 100 Hz

TS4 Température °C 1 Hz

PS6 Pression bar 100 Hz

VS1 Vibration mm/s 1 Hz

EPS1 Puissance moteur W 100 Hz

CE Efficacité de refroidissement (virtuelle)

FS1 Débit l/min 10 Hz

% 1 Hz

Chapitre IV : application d'algorithme de fourmis et KNN pour la sélection et la classification des caractéristiques

CP Puissance frigorifique (virtuelle) kW 1 SE Facteur d'efficacité % 1 Hz
Hz

Les valeurs de condition cible sont annotées par cycle dans « profile.txt » (délimité par des tabulations). Comme précédemment, le numéro de ligne Représente le numéro du cycle. Les colonnes sont :

1 : État du refroidisseur / % :

3 : proche de l'échec total

20 : efficacité réduite

100 : pleine efficacité

2 : État de la valve / % :

100 : comportement de commutation optimal

90 : petit décalage

80 : décalage important

73 : proche de l'échec total

3 : Fuite interne de la pompe :

0 : pas de fuite

1 : faible fuite

2 : fuite sévère

3 : Fuite interne de la pompe :

0 : pas de fuite

1 : faible fuite

2 : fuite sévère

4 : Accumulateur hydraulique / barre :

130 : pression optimale

115 : pression légèrement réduite

Chapitre IV : application d'algorithme de fourmis et KNN pour la sélection et la classification des caractéristiques

100 : pression fortement réduite

90 : proche de l'échec total

5 : la stabilité de système globale :

0 : les conditions étaient stables

1 : les conditions statiques n'ont peut-être pas encore été atteintes

IV.3. Conception et la méthodologie du la technique utilisé

En a traité dans notre travail la classification de l'états de système globale (état stable ou pas stable) et la classification des dégradations de la pompe et les dégradation de l'accumulateur

IV.3.1 Les défaut traité dans notre travail

1 : défaut de la pompe :

1 : pas de fuite

2 : faible fuite

3 : fuite sévère

2 : défaut de l'accumulateur :

1 : pression optimale

2 : pression légèrement réduite

3 : pression fortement réduite

4 : proche de l'échec total.

3 : la stabilité de système globale :

1 : état stables

2 : état conditions ne sont pas stable

Chapitre IV : application d'algorithmes de fourmis et KNN pour la sélection et la classification des caractéristiques

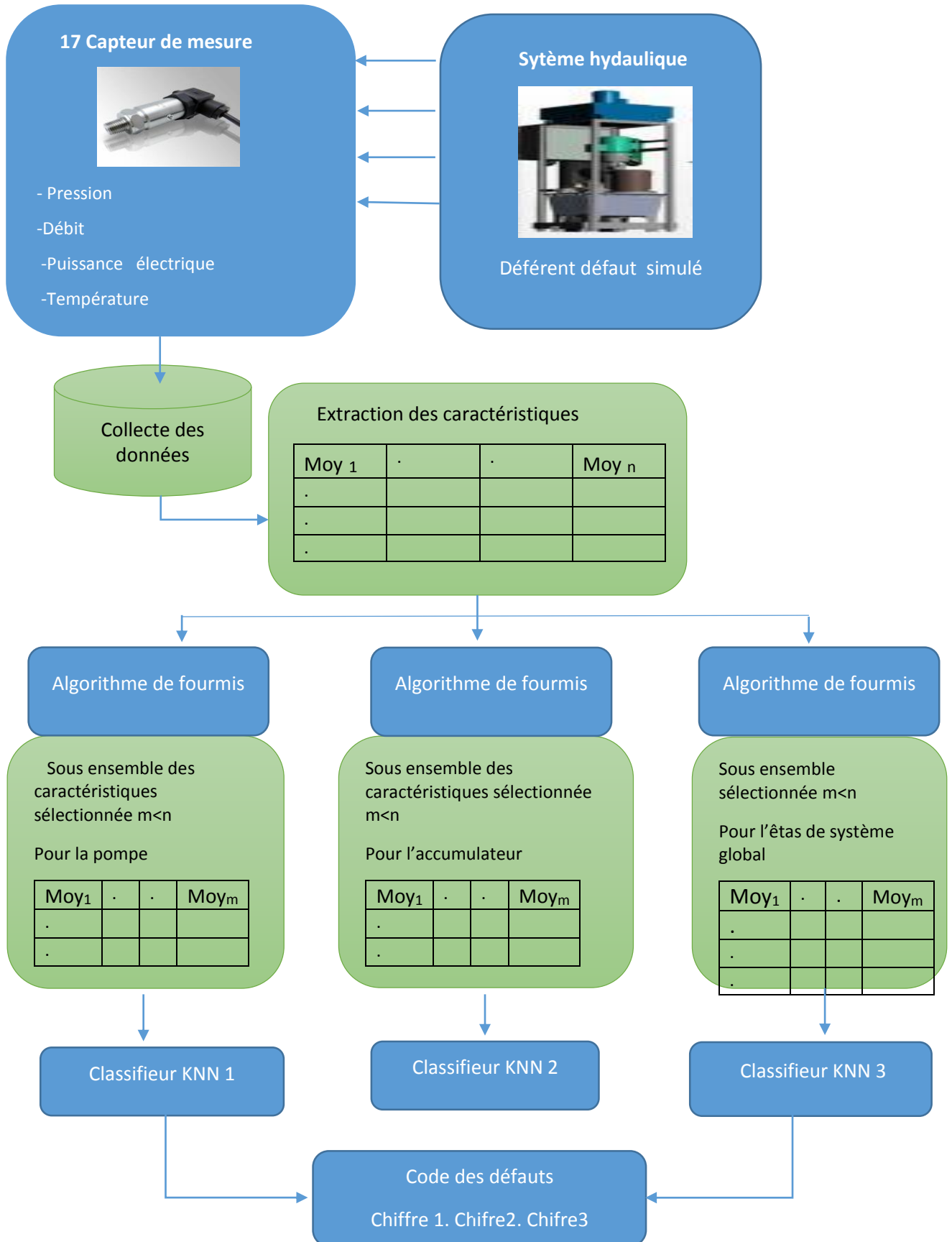


Figure IV.2 : organigramme de la méthode utilisée

Chapitre IV : application d'algorithme de fourmis et KNN pour la sélection et la classification des caractéristiques

La plus part des algorithmes de la classification fait l'apprentissage avec un seul vecteur de sortie (label) sauf quelque algorithme comme les réseaux de neurone et la logique floue, au début on a essayé avec le réseau de neurone artificiel, mais il n'a pas atteint l'objectif souhaité, il ne pouvait pas former un modèle (il ne comprenait pas la relation entre les entrées et les sorties).

Ensuite, nous avons pensé à utiliser la logique floue après avoir obtenu les règles à l'aide de l'arbre de décision, Les résultats d'arbre de décision obtenus ne suffisent pas à leur faire confiance pour créer les règles que nous utilisons pour l'apprentissage par l'algorithme de la logique floue. Donc nous avons choisi une méthode de création de trois classifieurs en parallèle (classifieurs pour chaque défaut) à la fin un code des défauts sera affichée (défaut de pompe, défaut de l'accumulateur, et l'état globale de système) qui nous a permis de donner des informations sur le système est- ce qu'il est stable ou instable et les type des défauts s'il est existé.

Nous avons essayé cette méthode par les algorithme SVM, l'arbre de décision, la forêt aléatoire et KNN, que nous avons choisi comme technique idéale l'algorithme KNN en raison de ses meilleurs résultats et de leur capacité à atteindre nos objectifs.

IV.3.2. Extraction des caractéristiques

Les moyenne de chaque capteur sera mesuré et combinées dans une matrice composée de 17 colonne représenté les capteur et 2205 ligne représenté le nombre de cycle d'essai.

IV.3.3. La sélection et la classification des caractéristiques

L'algorithme des fourmis sélectionne les caractéristiques de façon aléatoire et Nous pouvons contrôler le nombre des caractéristiques sélectionnée pour chaque défaut (on a choisi trois pour chaque défaut), puis l'algorithme de KNN classe ces caractéristiques et calculé le performance la classification (accuracy) pour le sous ensemble sélectionnée, L'algorithme effectue cette opération en fonction du nombre d'itérations que nous avons ajusté (T=100). Puis il choisit les caractéristiques qui nous ont donné le meilleur accuracy comme résultat final pour validée le sous ensemble optimale pour faire l'apprentissage.

Chapitre IV : application d'algorithme de fourmis et KNN pour la sélection et la classification des caractéristiques

A la fin du processus de la sélection, nous obtenons une matrice de trois colonnes pour chaque défaut, et nous pouvons voir de quels capteurs nous pouvons-nous passer et les quels le choisi.

Le programme effectue ce travail en trois phases :

La première phase pour la pompe et une fois terminée, elle va travailler les mêmes étapes sur l'accumulateur et enfin travailler sur la stabilité de système global.

IV.3.4. Evaluation et interprétation des résultats obtenue

Dans notre travail l'évaluation des résultats sont obtenus par la méthode hold-out

IV..3.4.1. Les résultats de la sélection

Après optimisation, le nombre de capteurs est devenu de 7 capteur :

1: Ps1 Pression bar 100 Hz

2: Ps2 Pression bar 100 Hz

3: Ps3 Pression bar 100 Hz

4: Ps5 Pression bar 100 Hz

5 : Fs2 Débit l/min 10 Hz

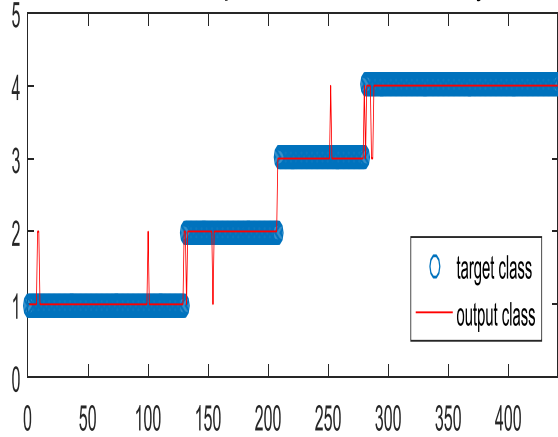
6 : CP Puissance frigorifique (virtuelle) kW 1 Hz

7 : Fs1 Débit l/min 10 Hz

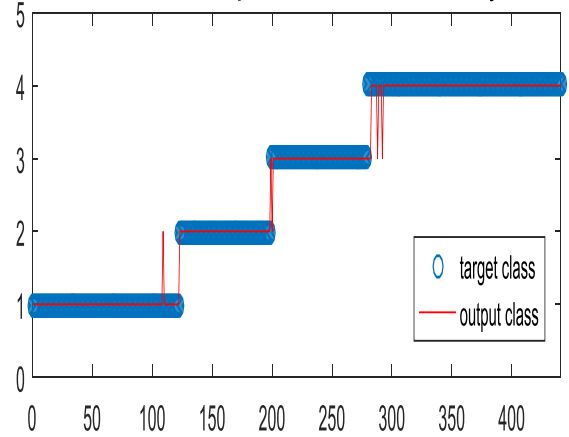
Chapitre IV : application d'algorithme de fourmis et KNN pour la sélection et la classification des caractéristiques

IV..3.4.2. Résultat de la classification

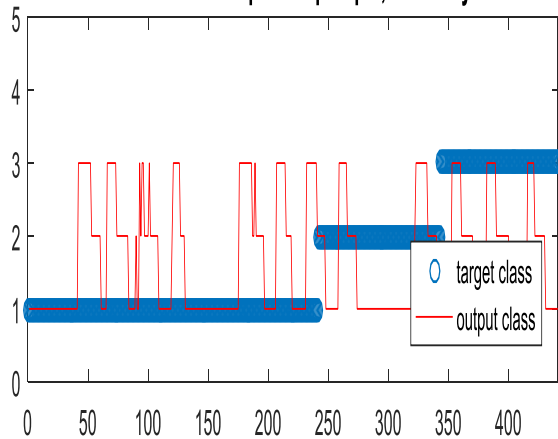
DT-classification results pour accumulateur , accuracy=97.2789%



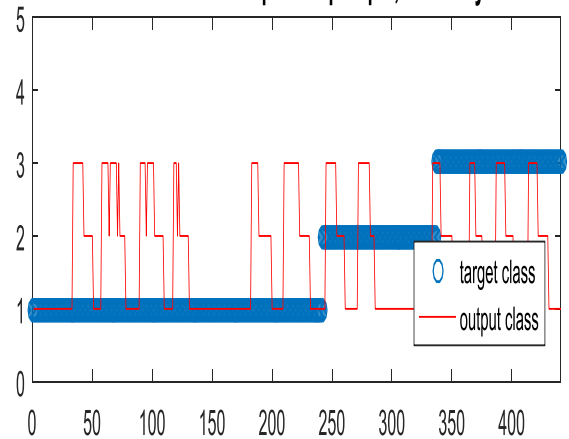
knn-classification results pour accumulateur , accuracy=98.4127%



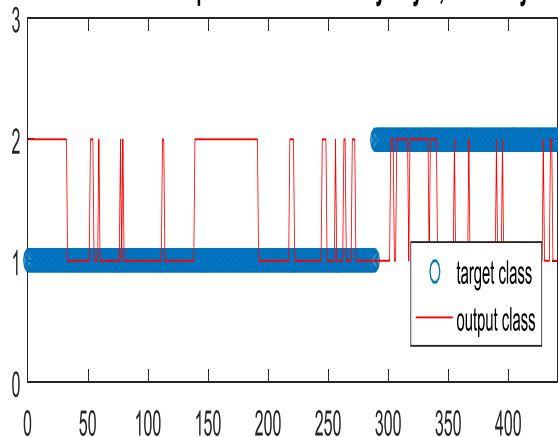
DT-classification results pour la pompe , accuracy=98.6395%



knn-classification results pour la pompe , accuracy=99.093%



DT-classification results pour la stabilité de sys hyd , accuracy=96.8254%



knn-classification results pour stabilité de syste hyd , accuracy=97.9592%

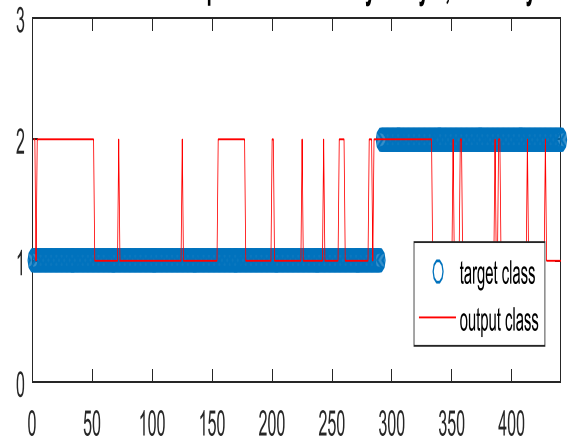


Figure IV.3: graphes des résultat accuracy

Chapitre IV : application d'algorithmes de fourmis et KNN pour la sélection et la classification des caractéristiques

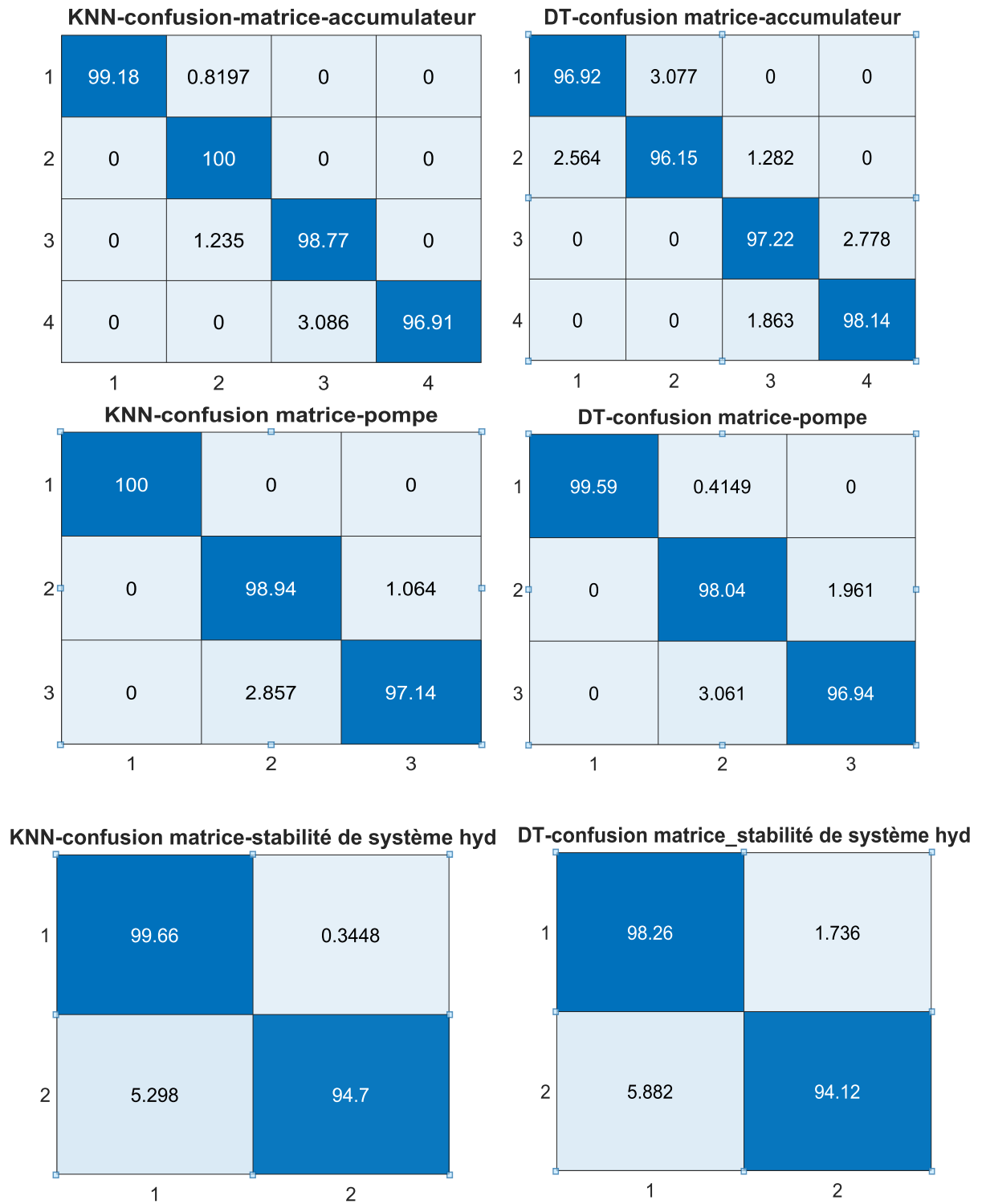


Figure IV.4 : les matrices de confusion

Chapitre IV : application d'algorithmes de fourmis et KNN pour la sélection et la classification des caractéristiques

La figure IV.3, représente les résultats de classification des deux techniques de classification pour les trois états de fonctionnement des systèmes hydrauliques.

La figure IV.4 représente les résultats de classification de chaque classe pour les trois états de fonctionnement des systèmes hydrauliques.

Les résultats obtenus montrent que le KNN présente une bonne performance, cependant cette figure ne donne aucune information sur la performance de classification de chaque classe. Pour cela nous avons utilisé les matrices de confusion représentées dans la figure 3. D'après cette figure on peut conclure que le KNN est le mieux placé pour la classification des défauts du système hydraulique.

Il ne reste d'étudier la stabilité pour cela nous avons réalisé 10 tests et nous avons ensuite calculé l'accuracy moyenne, le maximum, le minimum, et l'écart type. Les résultats obtenus sont représentés sur les tableaux IV.1, IV.2. Les résultats obtenus montrent que le KNN est le plus stable

	Accuracy accumulateur	Accuracy pompe	Accuracy de la stabilité de système hyd
	97,2789	98,6395	96,8254
	94,3311	99,3197	97,2789
	95,6916	98,4127	95,0113
	96,8254	99,093	96,3719
	97,2789	97,7324	94,3311
	96,8254	97,7324	96,1451
	96,8254	98,8662	97,0522
	97,0522	99,093	96,8254
	96,8254	98,4127	97,2789
	95,6916	97,7324	96,1451
Max	97,2789	99,3197	97,2789
Min	94,3311	97,7324	94,3311
Moyenne	96,46259	98,5034	96,32653
Ecart type	0,939207	0,606587	0,978537

Tableau IV.2 : résultat de classification par l'arbre de décision DT

Chapitre IV : application d'algorithmes de fourmis et KNN pour la sélection et la classification des caractéristiques

	Accuracy accumulateur	Accuracy pompe	Accuracy de la stabilité de système hyd
	98,4127	99,093	97,9592
	97,5057	99,093	97,2789
	97,0522	98,8662	97,7324
	97,2789	98,4127	98,1859
	97,2789	99,093	97,0522
	96,8254	98,1859	97,5057
	97,2789	98,4127	96,8254
	97,9592	97,2789	97,7324
	97,5057	99,5465	97,9592
	97,0522	97,9592	97,2789
Max	98,4127	99,5465	98,1859
Min	96,8254	97,2789	96,8254
Moyenne	97,41498	98,59411	97,55102
Ecart type	0,468385	0,67438	0,438126

Tableau IV.3: résultats de la classification par KNN avec k=3

Une fois le KNN a été selectionne nous procédons au choix de la valeur optimale de k

Pour cela nous avons réalisé différent teste pour différentes valeurs de k et pour les trois états de fonctionnements

Les résultats son Représenté dans les tableaux **IV 3, IV 4, IV 5**

Chapitre IV : application d'algorithmes de fourmis et KNN pour la sélection et la classification des caractéristiques

	Accuracy accumulateur	Accuracy pompe	Accuracy de l'état de système hyd
	96,3719	98,8662	97,9592
	97,2789	97,9592	98,1859
	96,8254	98,6395	97,5057
	96,5989	99,5465	97,9592
	97,0522	97,5057	97,9592
	96,8254	98,4127	96,8254
	96,5989	99,5465	99,093
	96,3719	99,3197	97,9592
	97,5057	98,6395	96,1451
	97,2789	98,4127	97,7324
Max	97,5057	99,5465	99,093
Min	96,3719	97,5057	96,1451
Moyenne	97,2789	98,4127	97,7324
Ecart type	0,39705571	0,66583372	0,79275965

Tableau IV.3: résultats de la classification par KNN avec k=5

Chapitre IV : application d'algorithmes de fourmis et KNN pour la sélection et la classification des caractéristiques

	Accuracy accumulateur	Accuracy pompe	Accuracy l'état de système hyd
	96,1451	98,8662	96,8254
	95,6916	97,9592	97,7324
	96,1451	98,4127	98,1851
	94,3311	97,9592	96,8254
	95,6916	98,4127	97,7324
	95,6916	97,5057	97,2789
	96,1451	98,4127	96,5986
	97,0522	98,6395	96,8254
	95,0113	97,5057	97,2789
	96,1451	97,5057	96,8254
Max	97,0522	98,8662	98,1851
Min	94,3311	97,5057	96,5986
Moyenne	96,1451	97,5057	96,8254
Ecart type	0,73477884	0,50193829	0,52404308

Tableau IV. 4: résultats de la classification par KNN avec k=7

Chapitre IV : application d'algorithmes de fourmis et KNN pour la sélection et la classification des caractéristiques

	Accuracy accumulateur	Accuracy pompe	Accuracy état de Sys hyd
	96,3719	99,093	95,9184
	97,7324	99,3197	95,9184
	97,7324	99,093	96,3719
	97,2789	98,6395	96,3719
	97,7324	98,6395	96,8254
	97,7324	98,6395	96,8254
	97,2789	99,093	96,3719
	96,3719	98,4127	97,7324
	99,5465	99,5465	96,7319
	97,2789	99,3197	97,2789
Max	99,5465	99,5465	97,7324
Min	96,3719	98,4127	95,9184
Moyenne	97,2789	99,3197	97,2789
Ecart type	0,88792866	0,37411673	0,57110834

Tableau IV. 5: résultats de la classification par KNN avec k=9

A travers ces résultats, nous pouvons voir l'effet de la valeur de K sur la performance,

Donc en besoin de choisir pour chaque élément de système hydraulique la valeur de K optimal du classifieur pour obtenir une bonne performance et une bonne stabilité.

- K pompe= 9
- K accumulateur= 5
- K pour tout le système= 3

IV.4. Conclusion

Dans ce chapitre, nous avons proposé une nouvelle méthode de diagnostic des dysfonctionnements des systèmes hydrauliques Sur la base d'un algorithme d'optimisation (algorithme de fourmi) et d'un algorithme d'apprentissage automatique (Classifieur) KNN, cette technique nous donné des bons résultats, donc nous pouvons faire confiance pour l'appliqué pour le diagnostic les systèmes hydrauliques

Conclusion générale

Les systèmes hydrauliques constituent dans nos jours une révolution industrielle du fait de son rôle essentiel dans de nombreux domaines,

La défaillance des systèmes hydraulique ou d'un de ses composants peut affecter les autres systèmes qui lui sont connectés

Pour cette raison on a proposé et travaillé sur une technique très efficace basée sur l'intelligence artificielle pour atteindre l'objectif de diagnostic des systèmes hydraulique.

On a fait dans notre travail le diagnostic d'un système hydraulique complexe contenant de grand nombre de capteur (17 capteur), et nous avons surmonté le problème du nombre de capteur par l'utilisation de la technique d'optimisation par algorithme des fourmis et on a minimiser le nombre des capteurs qui devient de 7 capteur ap

On a essayé avec la méthode de KNN autres méthodes, (SVM, Forêt aléatoire, et Arbre de décision). D'après la comparaison des résultats obtenus on a démontré l'efficacité de la méthode KNN par rapport à autre méthode.

Dans notre travail nous pouvions traiter seulement, les défauts de pompe, les défauts de l'accumulateur et la stabilité de l'état global de système, et nous ne pouvions pas traiter tous les défauts de manière simultanée.

Donc les perspectives sont ouvertes, pour étudier et proposer des méthodes pour diagnostiquer tous les défauts de manière simultanée.

Bibliographié :

- [1] **Djihed Bensaad**,(diagnostique de fuite interne dans une pompe à piston) ,THESE
THESE de DOCTORAT DE L'UNIVERSITE DE LYON,03/09/2019.
- [2] **Ouhadj mahmoud**, (Etude technologie et analyse du rendement d'une pompe centrifuge ,Univ-Abderrahmane mira Bejaia ,faculté de technologie,12/10/2020
- [3] **Messaoui mosbah** (support de cours système hydraulique) ,institut supérieure des Études technologie 2010-2011 de Gafsa
- [4] **Sami Bellalah** (chapitre 2 :l'hydraulique industriel)
- [5] **Thierry Boulay** (hydraulique industrielle)
- [6] **Mouhamed walid Azizi** (cours hydraulique et pneumatique)
- [7] **A.Akhenak**. Conception d'observateurs non linéaires par approche multi modèle, application au Diagnostic, Thèse de Doctorat, Ecole doctorale IAEM de Lorraine, 2004
- [8] **Said Alem** ,(cours maintenance) univ-boumerdes
- [9] **R.J.Pattan**. Fault Detection and Diagnosis in Aerospace Systems using Analytical Redundancy, Computing and Control Engineering, Vol.2, pp127-136, 1991.
- [11] **Boughagha Fouzi**. Diagnostic et détection des défauts mécaniques affectant les systèmes électromécaniques. UNIVERSITÉ BADJI MOKHTAR- ANNABA.2019
- [12] **Bensana Toufik** . (Diagnostic des défaillances basé sur l'analyse vibratoire d'une turbine à vapeur) . UNIVERSITE BADJI MOKHTAR – ANNABA.2008
- [12] **Microsoft Expériences**, Tout savoir sur l'Intelligence Artificielle, (consulté le 1/06/2021), disponible sur : <https://experiences.microsoft.fr/business/intelligenceartificielle-ia-business/comprendre-utiliser-intelligence-artificielle/>
- [13] **Bouguessa, M.** (2015b). Forage de données. notes de cours. Université du Québec à Montréal , 88 pages
- [14]. **J. Ah-Pine (Univ-Lyon 2)** Apprentissage automatique M2 DM 2019/2020 8 / 359
- [15]. **Lo, A. W., & MacKinlay, A. C.** (2002). *A non-random walk down Wall Street*. Princeton University Press
- [16] **rachide Mifdal** ,le 12 novembre 2019, (Application des techniques d'apprentissage pour la prédiction de la tendance des titres financiers) , école supérieure université du Québec
- [17]. **Boudheb Tarik**. 2018/2019.privacy Preserving Classification of Biomedical Data. Université Djillali Liabes faculté des sciences exactes Sidi Bel Abbès. Page 32

- [18] **WendBenedo Arnaud Bienvenue ZOUNGRANA.**2020. (Application des algorithmes d'apprentissage automatique pour la détection de défauts de roulements sur les machines tournantes dans le cadre de l'Industrie 4.0), page 60
- [19] **Carlos Gershenson.** Artificial neural networks for beginners.
- [20]. **JOSE ALEJANDRO LOPEZ GONZALEZ.** Juin 2016.(exploitation des arbre de décision et des support vecteur machine en vue d'application dans l'analyse de texte université du Québec
- [21] **Fatima mahmoude .**mars 2018,(approche d'apprentissage automatique pour la détection du spam wep informatique :exploitation de divers caractéristique. Université de Québec à montérial.page 72
- [22] **JOSE ALEJANDRO LOPEZ GONZALEZ.** Juin 2016,exploitation de décision et des support vecteur machine en vue d'application dans l'analyse de texte Université de Québec à montérial.page 9
- [23] **WendBenedo Arnaud Bienvenue ZOUNGRANA.2020.** Mémoire présenté à l'Université du Québec à Chicoutimi en vue de l'obtention du grade de Maître ès sciences appliquée en ingénierie, profil recherche. (Application des algorithmes d'apprentissage automatique pour la détection de défauts de roulements sur les machines tournantes dans le cadre de l'Industrie 4.0) .page 43
- [24] **Mr MENASSEL Rafik .**2018 ;(Conception de métaheuristique d'optimisation pour la compression d'image) .Thèse doctorat.univ-BADJI MOKHTAR-ANNABA.page 23
- [25]. **Omran M. G. H.** "Particle Swarm Optimization Methods for Pattern Recognition and Image Processing", PHD Thesis, University of Pretoria, Pretoria, South Africa, 2004.
- [26] **Melle MENGHOUR Kamilia .** Année 2014-2015. (Approches Bio-inspirées pour la Sélection d'Attributs). THESE Présentée en vue de l'obtention du diplôme de Doctorat 3ème Cycle
- [27] **Pudil, P., Novovičová, J., et Kittler.,J., 1994.** "Floating search methods in feature selection". Pattern Recognition Letters, Vol.15, pp 1119–1125
- [28] **Jain, A. et Zongker, D.,1997.** "Feature selection: Evaluation, application and small sample performance". IEEE Trans. on Pattern Analysis and Machine Recognition, vol.19, n^o.2, pp. 153-158.
- [29], **Jain, A.K. Duin, R.P.W. et Mao. J., 2000.** " Statistical pattern recognition. A review", IEEETrans. on Pattern Analysis and Machine Recognition, vol.22, n^o.1, pp. 4-37.
- [30] **Chatelain, C., 2006.** " Extraction de séquences numériques dans des documents manuscrits quelconques ". Thèse de doctorat en informatique, Université de Rouen, 198 pages.
- [31] **Ishak, A. B. (2007).**Sélection des variables par la machine à vecteurs support pour la la discrimination binaire et multiclasse en grande dimension. dimension.
- [32] **Dash, M. and Liu, H. (1997).** Feature selection for classifica-tion, in intelligent data analysis.Intelligent *Data Analysis*, 1 :131–156.

- [33] **H. Liu, L. Yu**, "Toward integrating feature selection algorithms for classification and clustering," *Knowl. Data Eng. IEEE Trans. On*, vol. 17, no. 4, pp. 491–502, 2005.
- [34] **Blum, A. et Rivest, R. (1993)**. Training a 3-node neural network is np-complete. In Hanson, S.,
- [35] **Hassan CHOUAIB** . 8 juillet 2011, Sélection de caractéristiques : méthodes et applications. Thèse présentée pour l’obtention du grade de Docteur de l’université Paris Descartes.
- [36] **John, G. H., Kohavi, R. et Pfleger, K. (1994)**. Irrelevant features and the subset selection problem. In *MACHINE LEARNING : PROCEEDINGS OF THE ELEVENTH INTERNATIONAL*, pages 121–129. Morgan Kaufmann.
- [37] **R. Kohavi, G. H. John**, "Wrappers for feature subset selection,"
- [38] **Sabra Elferchichi** .(sélection et extraction d’attribut pour le problème de classification).Thèse doctorat .université de science et de technologie Lille 1
- [39] **M. Dorigo, V. Maniezzo, and A. Colorni**. The Ant System : Optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics – Part B*, 26(1) :29–41, 1996.
- [40] **M. Dorigo**. Optimization, Learning and Natural Algorithms (in Italian). PhD thesis, Dipartimento di Elettronica, Politecnico di Milano, Italy, 1992.
- [41] **Blum, A.L. and Langley, P. (1997)**, Selection of Relevant Features and Examples in Machine Learning. *Artificial Intelligence*
- [42] **Nikolai Helwig, Eliseo Pignanelli**, Condition Monitoring of a Complex Hydraulic System Using Multivariate Statistic, in *Proc. I2MTC-2015 - 2015 IEEE International Instrumentation and Measurement Technology Conference*, paper PPS1-39, Pisa, Italy, May 11-14, 2015, doi: 10.1109/I2MTC.2015.7151267.
- [43] **N. Helwig, A.** Detecting and compensating sensor faults in a hydraulic condition monitoring system, in *Proc. SENSOR 2015 - 17th International Conference on Sensors and Measurement Technology*, oral presentation D8.1, Nuremberg, Germany, May 19-21, 2015, doi: 10.5162/sensor2015/D8.1.