

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

UNIVERSITE FERHAT ABBAS – SETIF
UFAS (ALGERIE)

MEMOIRE

Présenté à la faculté des sciences de l'ingénieur
Département d'optique et mécanique de précision
Pour l'obtention du diplôme de

MAGISTER

Option : Optique et mécanique de précision

Par

Mr. FENINECHE HOCINE

THEME

**Application des réseaux de neurones artificiels au
diagnostic des défauts des machines tournantes**

Soutenu le :

devant la commission d'examen :

BOUZID Djamel	Prof	à l'université de Sétif	Président
ZEGADI Rabah	Prof	à l'université de Sétif	Rapporteur
FELKAOUI Ahmed	M.C	à l'université de Sétif	Co-Rapporteur
CHIKOUCHE Djamel	Prof	à l'université de M'sila	Examineur
BOUZID Said	M.C	à l'université de Sétif	Examineur

Résumé :

Les réseaux de neurones artificiels par leur capacité d'apprentissage, de classification, et de décision, ont contribué dans le développement de plusieurs domaines. En maintenance industrielle, les réseaux de neurones sont utilisés pour résoudre les problèmes du diagnostic par la classification automatique des signaux vibratoires correspondant aux différents modes de fonctionnement des machines. Ce travail porte sur l'application des réseaux de neurones artificiels pour l'automatisation de diagnostic vibratoire des défauts des machines tournantes. Le réseau de neurones utilisé est un réseau de type perceptron multicouches, entraîné par l'algorithme de la rétropropagation du gradient. Le vecteur de formes contient des coefficients extraite des signaux prélevés à partir des bancs d'essai étudiés. Le vecteur de sortie contient les différentes classes correspondant aux différents modes de fonctionnement des dispositifs expérimentaux. Les performances du réseau ont été étudiées en variant les paramètres du réseau et la structure des vecteurs de sortie et d'entrée. Les résultats obtenus montrent que le prétraitement des données d'apprentissage et du test (entrée du réseau), le choix des classes (sortie du réseau), et le choix des paramètres (configuration) du réseau ont une importance majeure sur les performances du réseau de neurones artificiels utilisé comme outil de diagnostic des machines tournantes.

Mots clés :

Machines tournantes, Vibrations mécaniques, Réseaux de neurones artificiels, Diagnostic des défauts, Traitement du signal.

Dédicace

À la mémoire de mon grand-père,

À mes parents,

À mes frères et sœurs,

Et à toute la famille,

À mes deux petites nièces,

Nour-Elimane Et Inas,

À tous mes amis,

Je dédie ce travail.

Remerciements

Les travaux présentés dans ce mémoire ont été réalisés au Laboratoire de la Mécanique de Précision Appliquée (LMPA), du département d'Optique et Mécanique de Précision (UFAS – SETIF) au sein du groupe Diagnostic et Surveillance des Machines Tournantes.

Mes remerciements s'adressent tout d'abord à mes directeurs de thèse, Monsieur le professeur Zegadi Rabah, directeur du laboratoire LMPA, et Monsieur Felkaoui Ahmed, Docteur à DOMP-UFAS. Je les remercie pour leur disponibilité et leurs précieux conseils. Je voudrais les remercier aussi pour toute la confiance qu'ils ont eu en moi, pour leur patience et leur soutien.

Je suis très reconnaissant de l'honneur que m'a fait Monsieur Bouzid Djamel, Professeur à DOMP-UFAS, d'avoir accepté de présider ce jury. J'exprime également ma reconnaissance à Messieurs Chikouche Djamel, Professeur à l'université de M'sila, Bouzid Saïd, Maître de conférence à DOMP-UFAS, pour leur participation à ce jury.

Durant ces années passées au DOMP, j'ai pu apprécier l'amitié de certaines personnes qui, au-delà de l'ambiance très amicale qu'ils ont créée, m'ont toujours apporté leur soutien dans les moments critiques. Je voudrais remercier très sincèrement mes collègues, et l'ensemble du personnel de département d'optique et mécanique de précision.

Ma gratitude et mes très sincères remerciements s'adressent à ma famille, qui m'a toujours encouragé, et soutenu dans les moments difficiles.

Enfin, je tiens à exprimer ma gratitude envers toutes les personnes qui ont contribué scientifiquement et humainement à la réalisation de ce travail.

TABLE DES MATIERES

INTRODUCTION GENERALE.....	8
CHAPITRE 1 : DEFAUTS DES MACHINES TOURNANTES	
 ET TECHNIQUES D'ANALYSE VIBRATOIRE	
1.1. Introduction	11
1.2. Principaux défauts des machines tournantes.....	11
1.2.1 Défauts liés à l'arbre	11
a- Balourd.....	11
b. Désalignement.....	12
1.2.2. Défauts d'engrenages.....	13
a. Défauts uniformes.....	13
b. Défauts localisés.....	14
c. Vibrations d'engrenages.....	15
1.2.3. Défauts de roulements.....	16
a. Défaut sur la bague interne.....	16
b. Défaut sur bague externe.....	17
c. Défauts sur les éléments roulants.....	17
d. Défauts sur la cage.....	17
e. Les vibrations d'un roulement.....	17
1.3. Techniques d'analyse vibratoire.....	17
1.3.1. Techniques classiques.....	18
1.3.1.1. Méthodes temporelles.....	18
1.3.1.2. Méthodes fréquentielles.....	18
a. Analyse spectrale.....	19
b. Analyse cepstrale	19
1.3.1.3. Méthodes temps-fréquence et temps –échelle.....	19
a. Temps- fréquence.....	19
b. Temps- échelle.....	20
1.3.1.4. Analyse en enveloppe.....	21
1.3.1.5. Méthodes paramétriques.....	22
1.3.2. Méthodes automatisées.....	22
1. Système expert.....	22

2. Logique floue.....	23
3. Séparateurs à vaste marge (SVM).....	23
4. Analyse discriminante.....	24
5. Réseaux de neurones artificiels.....	24
1.4. Conclusion.....	26

**CHAPITRE2 : ETAT D'ART DE L'APPLICATION DES RESEAUX DE
NEURONES ARTIFICIELS (RNA) DANS LE DIAGNOSTIC DES MACHINES
TOURNANTES**

2.1. Introduction.....	29
2.2. Généralités sur les réseaux de neurones.....	29
2.2.1. Neurone biologique.....	29
2.2.2. Réseau formel.....	30
2.2.3. Apprentissage des Réseaux de neurones artificiels.....	31
2.2.4. Architecture des RNA.....	32
2.3- Bibliographie sur l'application des RNA dans le diagnostic des machines tournantes.....	33
2.4. Conclusion.....	44

CHAPITRE 3 : DEMARCHE EXPERIMENTALE

3.1. Introduction.....	47
3.2. Réseau perceptron multicouches.....	47
3.2.1. Neurones d'entrée.....	48
3.2.2. Neurones cachés.....	48
3.2.3. Noeuds de sortie.....	48
3.2.4. Apprentissage du réseau perceptron multicouche.....	49
3.3. Déroulement de la classification par PMC.....	50
3.3.1. Acquisition des signaux.....	51
3.3.2. Prétraitement des données.....	51
3.3.3. Construction du vecteur forme.....	51
3.3.4. Apprentissage et test du réseau.....	51
3.3.5. Exploitation du réseau.....	51

3.4. Implémentation du PMC sous MATLAB.....	51
3.4.1. Création du réseau.....	52
3.4.2. Apprentissage du réseau.....	52
3.4.3. Test et validation du réseau.....	53
3.5. Choix des paramètres du réseau	53
3.5.1. Type du réseau.....	53
3.5.2. Prétraitement des données.....	53
3.5.3. Nombre des couches cachées.....	54
3.5.4. Nombre des neurones.....	54
3.5.5. Initialisation des poids.....	54
3.5.6. Taux d'apprentissage	54
3.5.7. Fonction d'activation.....	55
3.5.8. Algorithme d'apprentissage.....	55
3.5.9. Critères d'arrêt.....	55
3.6. Description des bancs d'essai.....	56
3.6.1 Banc d'essai d'engrenage (CETIM).....	56
3.6.1.1. Rapport d'expertise.....	57
3.6.1.2 Représentation des signaux temporels.....	58
3.6.2. Banc d'essai de roulement.....	58
3.7. Prétraitement et extraction des données.....	61
3.8. Classification par RNAs.....	61
3.9. Conclusion.....	62

CHAPITRE : 4

RESULTATS ET DISCUSSION

4.1. Introduction.....	65
4.2. Diagnostic des défauts d'engrenage.....	65
4.2.1. Constitution de la base de données.....	65
4.2.1.1. Préparation des données.....	66
4.2.1.2. Prétraitement de données.....	67
a- Filtrage.....	67
b- Décimation.....	67
4.2.1.3. Calcul des paramètres AR.....	68
4.2.1.4. Constitution du vecteur forme.....	68

4.2.1.5. Choix des classes.....	68
4.2.1.6. Normalisation des données.....	69
4.2.2. Exploitation du système du diagnostic basé sur les RNA.....	69
4.2.2.1. Configuration du RNA.....	69
4.2.2.2. Effet du nombre d'itérations maximal.....	70
4.2.2.3. Effet du nombre des neurones dans la couche cachée.....	72
4.2.2.4. Effet de la valeur du taux d'apprentissage.....	74
4.2.2.5. Effet de la valeur de l'erreur minimale.....	76
4.2.2.6. Effet du nombre d'exemples d'apprentissage.....	78
4.3. Diagnostic des défauts de roulements.....	80
4.3.1. Constitution de la base de données.....	80
4.3.1.1. Prétraitement et calcul des indicateurs.....	82
4.3.1.2. Constitution du vecteur forme.....	82
4.3.1.3. Choix des classes.....	83
4.3.1.4. Normalisation des données.....	83
4.3.2. Configuration du RNA.....	83
4.3.3. Effet du nombre de neurones dans la couche cachée.....	84
4.3.4. Choix des valeurs optimales des critères d'arrêt.....	86
a- Effet du nombre d'itérations maximal.....	86
b- Effet de l'erreur minimale.....	88
4.3.5. Effet du taux d'apprentissage.....	89
4.4. Conclusion.....	90
 CONCLUSION GENERALE	 93
REFERENCES BIBLIOGRAPHIQUES.....	96
ANNEXES	

INTRODUCTION GENERALE

INTRODUCTION GENERALE

La maintenance préventive des machines est devenue ces dernières trentaines années, l'un des facteurs les plus importants dans plusieurs domaines (industrie, transport,...). Cette importance est due au rôle du diagnostic dans la réduction des coûts de pannes et des taux de défaillances, car les pannes imprévues sont quelques fois très coûteuses. La perte de production pendant les réparations engendre un manque à gagner qui peut affecter les bénéfices de l'entreprise, si l'on y ajoute les problèmes de sécurité insuffisante pour le personnel, la qualité de production amoindrie et la perte éventuelle de réputation pour l'entreprise. Mais une telle maintenance ne peut être efficace, si on ne prend pas en considération la méthode du diagnostic, la nature et la complexité de l'installation des éléments des machines surveillés, et la grandeur mesurée (la pression, la température, le courant, les vibrations, ...etc.).

De nombreuses méthodes de diagnostic ont ainsi été développées, tel que l'analyse vibratoire, l'analyse des lubrifiants, la thermographie et l'analyse acoustique. Les techniques actuelles de surveillance des machines sont basées sur l'analyse vibratoire en exploitant le signal fourni par les accéléromètres (fixés à proximité des machines). Le diagnostic par analyse vibratoire a une large utilisation dans la maintenance car les vibrations engendrées par une machine traduisent les efforts dynamiques engendrés par les éléments mécaniques en mouvements. L'analyse et le suivi dans le temps de cette signature vibratoire permettent de dresser un diagnostic de l'état de fonctionnement de la machine.

Cependant, il existe deux approches de diagnostic basé sur l'analyse vibratoire; l'une classiques et la seconde automatisée. L'approche classique est basée sur le traitement de signal vibratoire prélevé des machines. Dans ces méthodes, la décision de l'existence ou non d'un défauts est faite par l'homme. Cette décision reste toujours imprécise et elle est surtout influée par l'erreur humaine; ce qui influe directement sur les performances du diagnostic. Nous n'oublions pas ces limitations dans le cas des machines complexes (défauts combinés). L'approche automatisée

(moderne) est basée sur la classification automatique des différents états de fonctionnement de la machine. La décision de l'existence ou non d'un défaut se fait automatiquement, en se basant sur différents algorithmes et règles offertes par le développement des diverses disciplines, et en particulier celles de l'intelligence artificielle. Mais, ces méthodes nécessitent une base de données importante, pour qu'elles puissent donner de bons résultats.

Parmi ces méthodes automatisées, on trouve en premier lieu les réseaux de neurones artificiels. A l'origine, l'idée recherchée dans cette méthode est de remplacer le raisonnement humain dans la procédure d'analyse des données mesurées à partir des machines, et de résoudre les problèmes de classification et de décision. Aussi, le but est d'augmenter les performances et de réduire les coûts du diagnostic industriel.

Dans notre travail, nous essayons d'appliquer cette méthode dans le diagnostic et la détection des défauts dans les machines tournantes (le cas de roulements et d'engrenages). Nous avons utilisé le réseau perceptron multicouche. Une comparaison entre ses performances, en variant ses paramètres et en utilisant différentes méthodes de prétraitement, sera effectuée.

- Dans le premier chapitre, nous présentons les défauts les plus courants dans les machines tournantes et les différentes méthodes de diagnostic par analyse vibratoire.
- Le deuxième chapitre, comporte un aperçu théorique sur les réseaux de neurones artificiels, et une synthèse bibliographique sur leur utilisation dans le diagnostic des machines tournantes.
- Le troisième chapitre, porte sur la démarche expérimentale d'un système de diagnostic basé sur les réseaux de neurones artificiels, et l'implémentation de ces derniers sous Matlab, en discutant le choix des différents paramètres pertinent à la structure des réseaux de neurones. Aussi, nous présentons les deux bancs d'essais et les signaux vibratoires.
- Le quatrième chapitre, comporte les résultats expérimentaux et leurs interprétations.

Enfin, nous terminons par une conclusion générale, ainsi que quelques perspectives d'amélioration de ce présent travail.

CHAPITRE 1
DEFAUTS DES MACHINES
TOURNANTES ET TECHNIQUES
D'ANALYSE VIBRATOIRE

CHAPITRE 1 : DEFAUTS DES MACHINES TOURNANTES ET TECHNIQUES D'ANALYSE VIBRATOIRE

1.1- Introduction :

Toutes les machines vibrent, et les vibrations générées représentent la somme de toutes les forces qui s'exercent sur la machine; la vibration mesurée est appelée signature vibratoire. Une quelconque modification, dans cette signature, nous renseigne sur l'état de santé de la machine. Le suivi de l'évolution dans le temps de la signature vibratoire, par les techniques d'analyse et d'interprétation des mesures de vibrations, aide à la mise en place d'une maintenance prédictive.

Dans la première partie de ce chapitre, nous présentons les défauts les plus courants dans les machines tournantes. Nous exposons ensuite, dans la deuxième partie les méthodes utilisées dans le diagnostic de ces défauts par l'analyse des signaux vibratoires.

1.2- Principaux défauts des machines tournantes :

1.2.1- Défauts liés à l'arbre :

L'arbre est l'organe principal de la machine. C'est lui qui supporte et entraîne les organes mobiles en rotation, tels que les roues, les engrenages, etc. Le fonctionnement d'un arbre est caractérisé par un mouvement de rotation autour de son axe de rotation supporté par deux paliers à ses extrémités.

Nous présentons deux types de défauts pour les arbres, le déséquilibre massique des rotors (balourd) et la dégradation du lignage (désalignement) [NORF06].

a- Balourd :

Un balourd, représenté par la figure 1.1, est caractérisé par un déséquilibre massique autour de l'axe de rotation qui produit des forces d'inertie, ces forces agissent sur les corps en rotation et surtout sur le roulement et excitent les paliers.

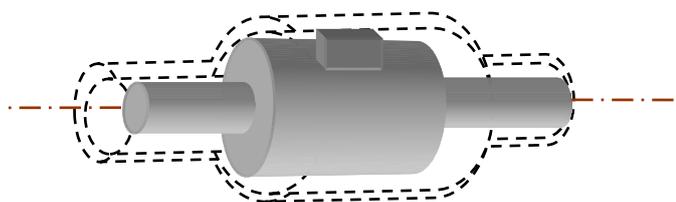


Fig1.1- Représentation du balourd.

L'allure spectrale, d'un balourd, est présentée par la figure 1.2. Elle est caractérisée par une composante dominante à la fréquence de rotation (f_r) de l'arbre.

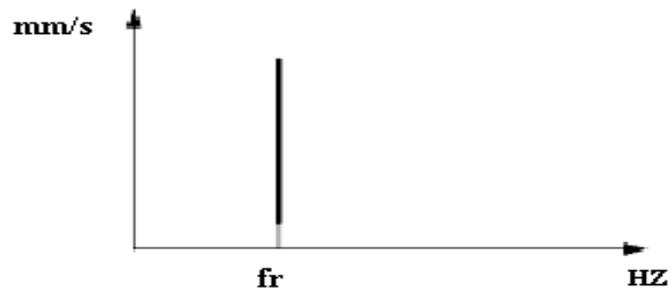
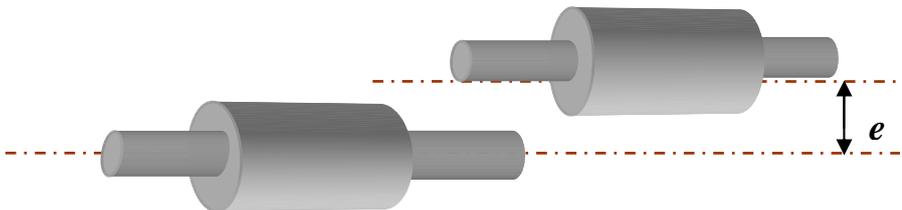


Fig1.2- Allure spectrale d'un balourd.

b. Désalignement :

Le désalignement est dû au fait que l'arbre menant et l'arbre mené n'ont pas le même axe de rotation. Dans la figure 1.3, nous présentons les différents types de désalignement; désalignement axial et angulaire.

- Désalignement axial:



- Désalignement angulaire :

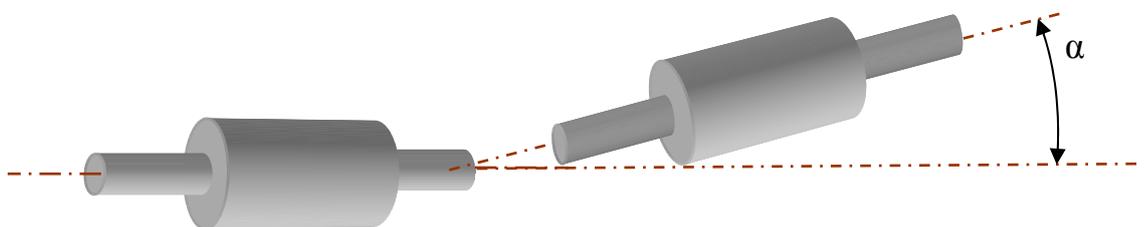


Fig1.3- Représentation des deux défauts de désalignement

Dans l'allure spectrale d'un défaut de désalignement, comme elle est présentée par la figure 1.4, la fréquence de rotation et les harmoniques 2 et 3 dominent le spectre dans les deux directions radiale et axiale.

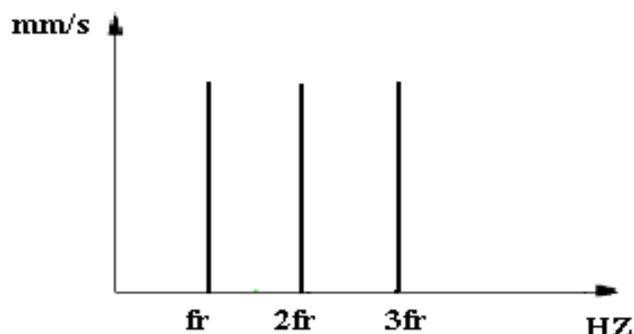


Fig1.4- Allure spectrale du désalignement.

1.2.2- Défaits d'engrenages :

Les engrenages ont pour fonction de transmettre une puissance d'un arbre en rotation à un autre arbre tournant à une vitesse généralement différente, les deux vitesses ont un rapport constant.

On peut distinguer pour les dentures d'engrenages deux types de défauts, des défauts uniformes et d'autres localisés [SIDAH90].

a- Défaits uniformes :

Les défauts uniformes concernent, en général toutes les dents, on peut en citer deux exemples :

L'usure, représentée par la figure 1.5, est un enlèvement de matière dû au glissement des deux surfaces de contact. Il peut modifier le profil des dents et conduire à un mauvais fonctionnement de l'engrènement.



Fig1.5- Défaut d'usure.

Le défaut de piqûres (pitting) est présenté par la figure 1.6. Il s'agit de trous plus ou moins profonds qui affectent toutes les dents, il peut évoluer suivant les matériaux jusqu'à la destruction de la denture.



Fig1.6- Défaut de piqûres.

b- Défauts localisés :

Quant aux défauts localisés, ils ne concernent que quelques dents et se répartissent en trois types : La fissuration, l'écaillage et la fracture.

- La fissuration d'une dent représentée par la figure 1.7 conduit rapidement à la rupture de la dent.



Fig1.7- Défaut de fissuration.

- L'écaillage est représenté par la figure 1.8. Il s'agit également de trous; mais ceux-ci sont moins nombreux, plus profonds et plus étendus que ceux du pitting.



Fig1.8- Défauts d'écaillage.

-

- La fracture qui est représentée par la figure 1.9, est un état de dégradation avancé de l'engrenage, elle génère des chocs importants sur le signal temporel.



Fig1.9- Défaits de fracture.

c- Vibrations d'engrenages :

Les vibrations induites par les engrenages ont une amplitude importante. Même en l'absence de défauts, les impulsions données à chaque engrènement provoquent des vibrations émergeant nettement du niveau global du bruit [LEBO00].

Le défaut soit significatif, si l'amplitude de la fréquence (raie) d'engrènement augmente ou s'il y a une génération d'autres fréquences (harmoniques et modulation de la fréquence d'engrènement) [THOM02]. Les figures 1.10, 1.11 et 1.12 représentent les allures spectrales des différents cas d'une denture d'engrenage.

• Denture correcte :

Une denture correcte montre les fréquences f_i et f_j des deux arbres plus la fréquence d'engrènement f_e .

La fréquence d'engrènement $f_e = N_i \cdot f_i = N_j \cdot f_j$

où N représente le nombre de dents.

• Une seule dent détériorée sur un pignon

- Si une seule dent est détériorée, il y aura un choc périodique à chaque tour de la roue.
- On verra donc de nombreuses harmoniques de la fréquence de rotation de l'engrenage f_i en cause ($n \cdot f_i$).
- La gravité du problème sera en fonction du nombre d'harmoniques (5 et plus).

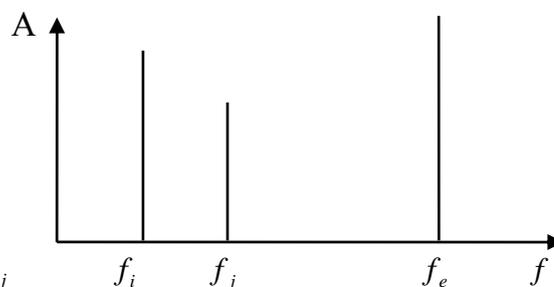


Fig1.10- Représentation fréquentielle d'une denture correcte

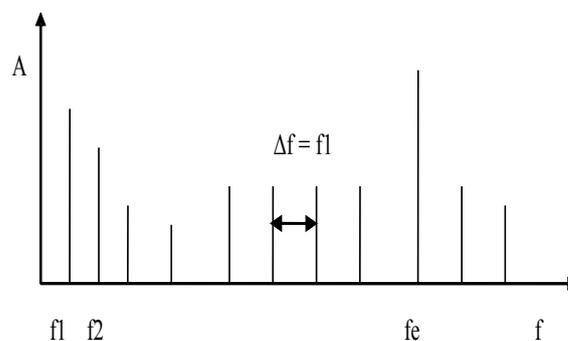


Fig1.11- Représentation fréquentielle d'une seule dent détériorée.

• Ensemble des dentures détériorées :

Si on détecte de nombreuses harmoniques de la fréquence d'engrènement f_e ($3f_e$ et plus un choc à chaque dent), Alors l'ensemble des dentures est détérioré.

- La gravité du problème n'est pas montrée par les amplitudes, mais par le nombre d'harmoniques (xf_e).

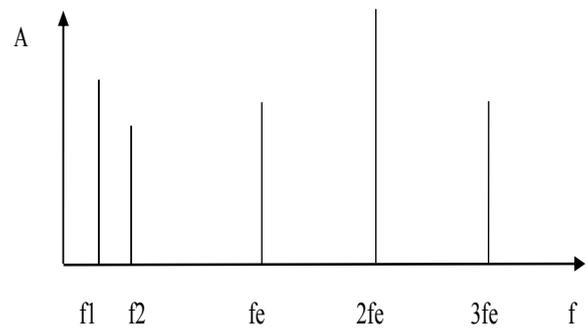


Fig1.12- Représentation fréquentielle d'un Ensemble des dentures détériorées

1.2.3- Défauts de roulements :

Un roulement est un organe qui assure une liaison mobile entre deux éléments d'un mécanisme roulant sur l'autre. Il permet leur rotation relative, sous charge, avec précision et frottement minimal.

Les défauts de roulements sont la conséquence d'un balourd, d'un désalignement, d'un problème d'engrenage ou d'un manque de lubrification [ROUG94].

Les vibrations générées par des défauts de roulement se manifestent à des fréquences particulières ainsi qu'à leurs harmoniques, qui en général ne sont pas synchrones avec la vitesse de l'arbre. Ces fréquences correspondent aux éléments d'un roulement [CARRE90].

La figure 1.13 représente les différents éléments d'un roulement.

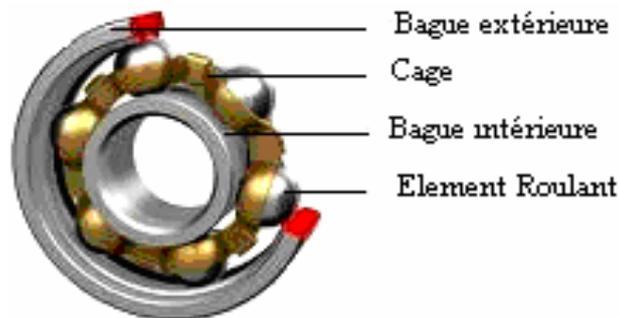


Fig1.13- Représentation d'un roulement

a- Défaut sur la bague interne :

Il est caractérisé par la présence d'une raie à la fréquence caractéristique du défaut f_{bi} .

$$f_{bi} = \frac{N}{2} \left(1 + \frac{d_b}{d_m} \cos(\alpha) \right) (f_{rot} - f_{ext}) \quad (1.1)$$

f_{rot} : Fréquence de rotation de la bague interne.

f_{ext} : Fréquence de rotation de la bague externe. En général elle est prise égale à zéro, si la bague extérieure est emmanchée (fixée) dans le carter.

N : Vitesse de rotation (rpm)

d_b : Diamètre des billes.

d_m : Diamètre moyen.

α : angle de contact.

b- Défaut sur bague externe :

Ce défaut est caractérisé par la présence d'un pic à la fréquence f_{be}

$$f_{be} = \frac{N}{2} \left(1 - \frac{d_b}{d_m} \cos(\alpha)\right) (f_{rot} - f_{ext}) \quad (1.2)$$

c- Défauts sur les éléments roulants :

La première fréquence caractéristique de défauts correspond à la fréquence de rotation de l'élément roulant f_{er} sur lui-même.

De plus, cet élément roulant rencontre une première fois la bague interne et une seconde fois la bague externe dans chaque tour. Alors, il génère des chocs à deux fois cette fréquence.

$$f_{er} = \frac{1}{2} \frac{d_m}{d_p} \left(1 - \frac{d_b}{d_m} \cos(\alpha)\right) (f_{rot} - f_{ext}) \quad (1.3)$$

d- Défauts sur la cage :

Ces défauts se manifestent par la présence d'un pic à la fréquence f_{ca} et ses harmoniques.

$$f_{ca} = \frac{1}{2} \left(1 - \frac{d_b}{d_m} \cos(\alpha)\right) f_{rot} + \frac{1}{2} \left(1 + \frac{d_b}{d_m} \cos(\alpha)\right) f_{ext} \quad (1.4)$$

e- Les vibrations d'un roulement :

Les vibrations générées au sein d'un roulement sont de faible amplitude, ressemblant à un bruit aléatoire. Lors de l'apparition d'une avarie, une impulsion se produit chaque fois que le défaut participe à un contact. L'avarie a donc une fréquence caractéristique qui dépend de la géométrie du roulement et de l'emplacement de l'écaillage (sur la bague intérieure, sur la bague extérieure, ou sur un élément roulant).

1.3- Techniques d'analyse vibratoire :

Le choix des méthodes de surveillance des machines tournantes est conditionné par la nature des installations et des dégradations éventuelles. Les experts de la maintenance n'ont jamais cessé depuis plusieurs années, de chercher les outils et les méthodes les plus performants de surveillance et de faire un diagnostic précoce des défauts des machines tournantes. Dans ce contexte, nombreuses études ont été faites. Néanmoins, la plupart d'entre elles ont été basées sur l'analyse des signatures vibratoires prélevées sur ces machines, en utilisant les techniques du traitement de signal; tel que l'analyse spectrale, les approches statistiques,...etc.

Quant à l'utilisation de ces méthodes dites classiques, la décision de la présence ou de l'absence des défauts reste toujours imprécise, surtout qu'elle est influée par l'erreur humaine, ce qui influe sur les performances du diagnostic. Ce problème pousse les experts de la maintenance de chercher à automatiser la décision, en profitant des nouvelles techniques et technologies, et surtout celles de l'intelligence artificielle.

1.3.1- Techniques classiques :

1.3.1.1- Méthodes temporelles :

L'analyse d'un signal vibratoire commence par la visualisation du signal temporel, comme le présente la figure 1.14, afin d'extraire certains paramètres : On peut citer les valeurs crête, valeur moyenne et surtout la valeur RMS.

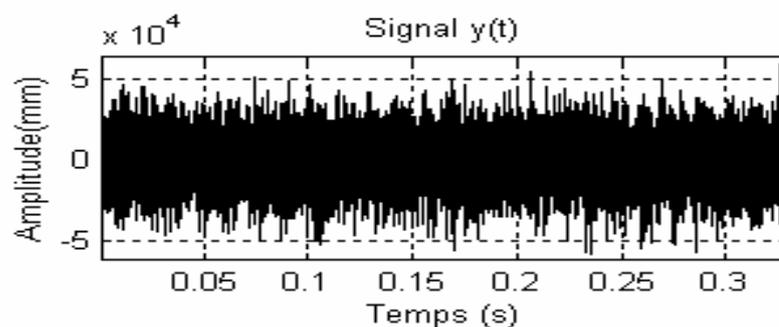


Fig1.14- Signal temporel.

En effet la valeur RMS prend en compte la forme de ce signal à cause de sa relation avec la puissance contenue dans le signal.

En outre, les paramètres (moments) statistiques servent à caractériser la distribution statistique des amplitudes d'un signal temporel vibratoire. On peut citer des indicateurs tel que la moyenne, l'écart type ou la variance, ou les moments d'ordre supérieur.

Parmi ces moments, le moment centré réduit d'ordre 4 appelé coefficient d'aplatissement (Kurtosis), ce moment est utilisé pour détecter la présence des chocs [BENOU99] .

1.3.1.2- Méthodes fréquentielles :

Il existe tout d'abord des méthodes basées sur le calcul de la transformée de Fourier rapide (FFT). Ce calcul est d'ailleurs accessible au niveau de n'importe quel analyseur portatif et souvent en temps réel. On trouve :

a- Analyse spectrale :

La densité spectrale de puissance (DSP) représentée par la figure 1.15, permet d'avoir une représentation du signal dans les domaines des fréquences. Elle permet de caractériser les phénomènes périodiques, par leurs fréquences et leurs amplitudes [ALATT94].

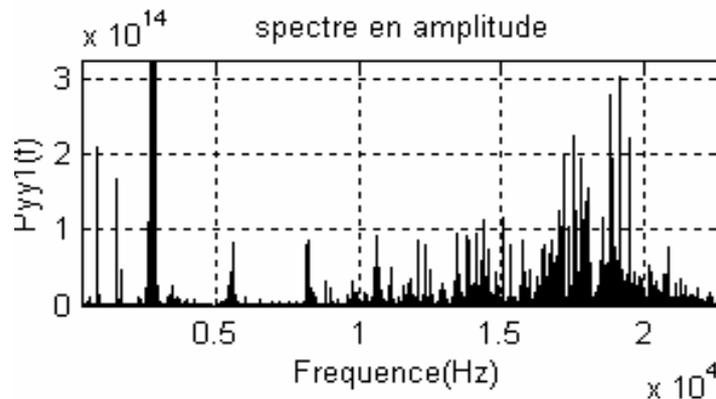


Fig1.15- spectre d'amplitude

b- Analyse cepstrale :

Le cepstre décrit sur la figure.1.16 est définie comme étant la transformée de Fourier inverse du logarithme de spectre de puissance. Il permet d'identifier la périodicité dans le spectre original, de quantifier et d'identifier dans le spectre des familles de composant et de suivre l'évolution dans le temps d'une ou plusieurs familles de défauts [ELBADA99a].

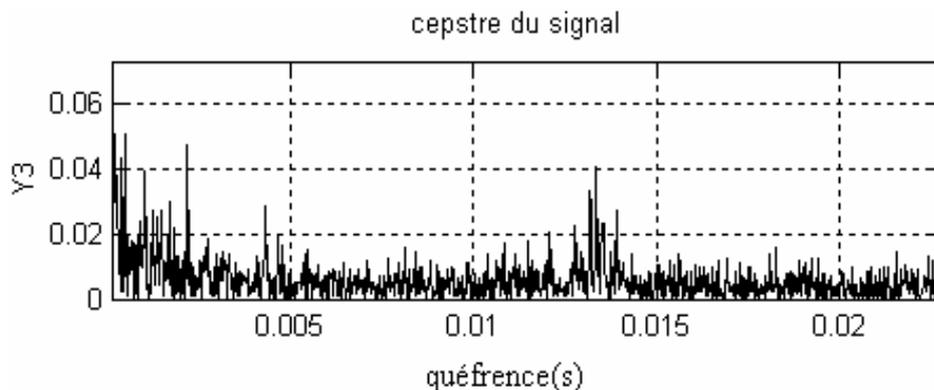


Fig1.16- Cepstre.

1.3.2.3- Méthodes temps-fréquence et temps –échelle :

a- Temps- fréquence :

Les méthodes d'analyse temps fréquence ont été développées pour l'étude des signaux non-stationnaires. Cette notion de non stationnarité peut s'appliquer aussi bien à des signaux aléatoires qu'à des signaux certains. Plus généralement, elles s'appliquent à des signaux dont le contenu fréquentiel ou les propriétés statistiques évoluent au cours du temps [BOUZ03]. Parmi ces méthodes on a la transformation de Wigner-Ville qui définit une distribution d'énergie dans le plan temps-fréquence, comme elle est décrite dans la figure 1.17.

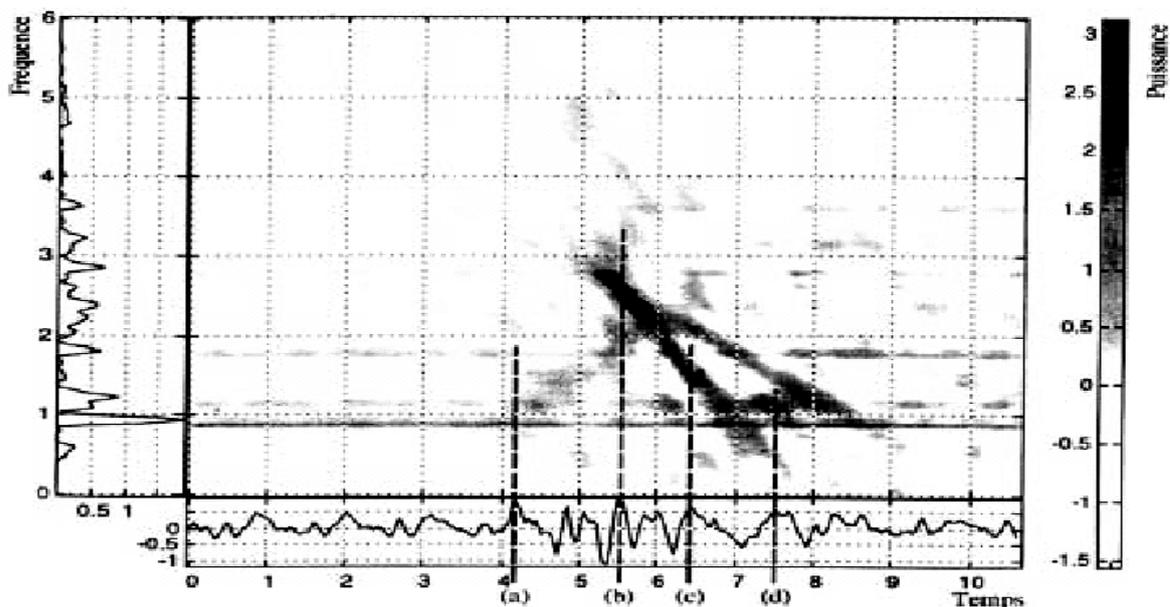


Fig1.17- Représentation temps- fréquence d'un signal vibratoire [OEHLM95].

b- Temps-échelle :

Cette méthode est utilisée pour détecter des phénomènes qui se déroulent sur des échelles de fréquence distinctes (figure 1.18), rencontrés dans un signal. La transformation en ondelette a été formalisée par Grossman et Morlet [STAR05]. Cette méthode consiste à décomposer le signal à l'aide de fonctions analysantes particulières construites à partir d'une ondelette «mère» oscillante et à moyenne nulle.

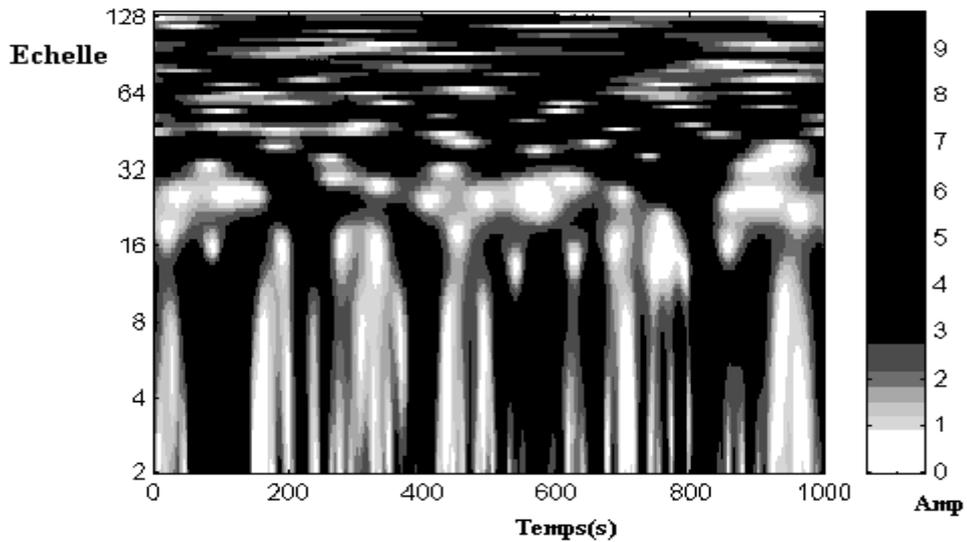


Fig1.18- Exemple d'une transformée d'ondelettes [LEPI01].

1.3.1.4- Analyse en enveloppe :

L'étude du signal d'enveloppe permet d'analyser la forme du signal. Elle consiste à filtrer le signal autour d'une fréquence de résonance identifiable sur le spectre. Cette technique est très utilisée dans le diagnostic des défauts de type choc, même pour de faibles vitesses, puisque les défauts ne sont pas recherchés dans la plage des basses fréquences du spectre mais autour de porteuses, dont les fréquences peuvent se situer à des centaines de fois la fréquence de rotation.

En pratique, la méthode de l'enveloppe demande une série de traitements du signal temporel brut avant d'obtenir le résultat. Ceux-ci sont repris sur la figure 1.19. [MILI06].

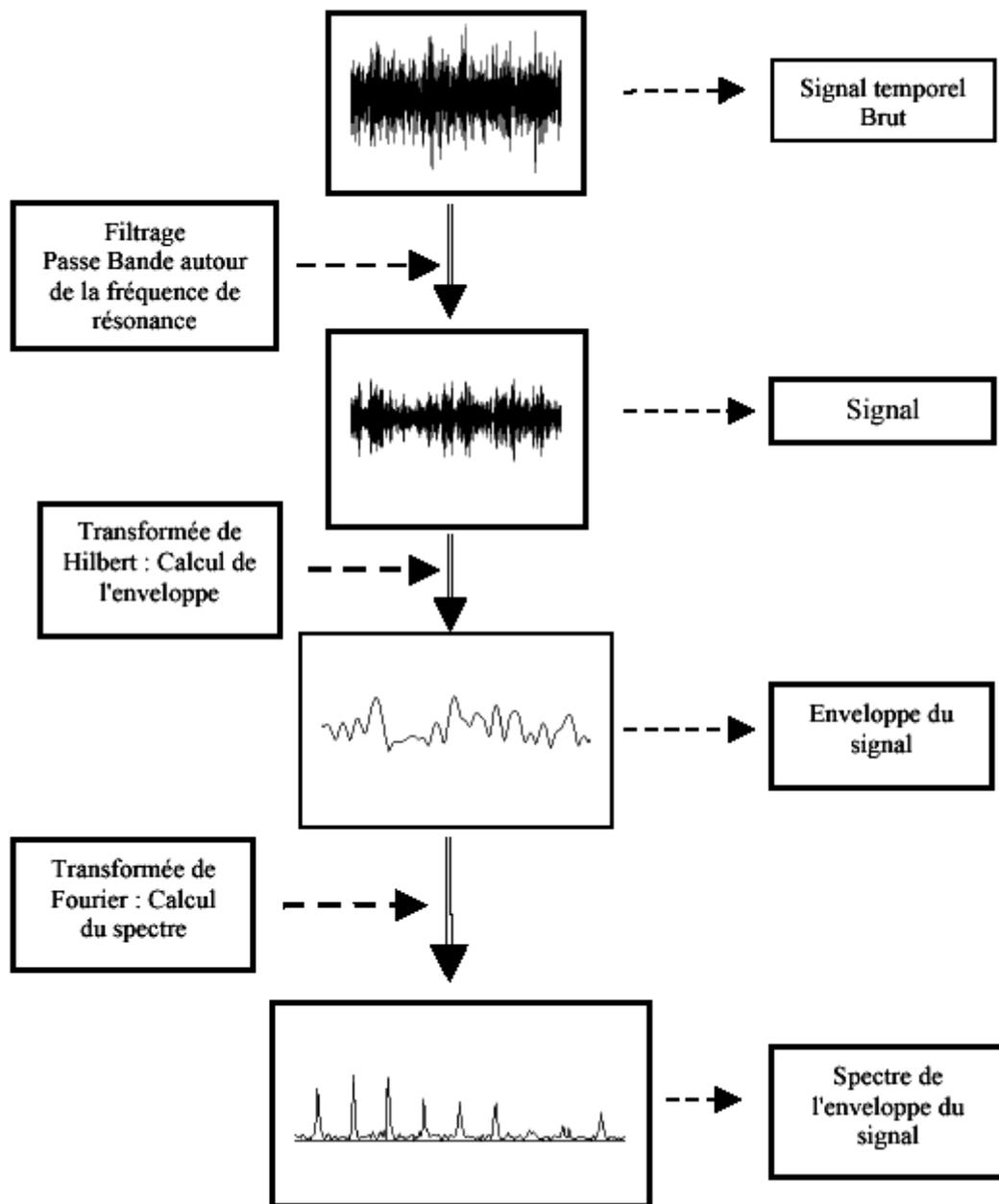


Fig1.19- Principe de calcul de l'enveloppe.

1.3.1.5- Méthodes paramétriques :

Ces méthodes consistent à appliquer des signaux de perturbation (bruit blanc) à l'entrée du système et en analyser la sortie dans le but d'obtenir un modèle purement mathématique. L'identification peut se faire soit dans le temps ou en fréquence. Les différents paramètres du modèle ne correspondent à aucune réalité physique dans ce cas. Ces paramètres peuvent être utilisés comme indicateurs de surveillance de l'état de santé de la machine. On a plusieurs modèles AR, ARMA, etc. [BROE06].

1.2.2- Méthodes automatisées :

1- Système expert :

Le système expert comme le souligne Gilles ZWINGELSTEIN dans [ZWIN95], est un système informatique destiné à résoudre un problème précis, à partir d'une analyse et d'une représentation des connaissances et de raisonnement d'un (ou plusieurs) spécialiste (s) de ce problème.

Un système expert est composé de deux parties indépendantes (voir Figure 1.20) :

- Une base de connaissances, est composée d'une base de règles qui modélise la connaissance du domaine considéré et d'une base de faits qui contient les informations concernant le cas traité.
- Un moteur d'inférences est capable de raisonner à partir des informations contenues dans l'espace de connaissances, afin de faire des déductions, etc. Au fur et à mesure que les règles sont appliquées des nouveaux faits se déduisent et se rajoutent à la base de faits.

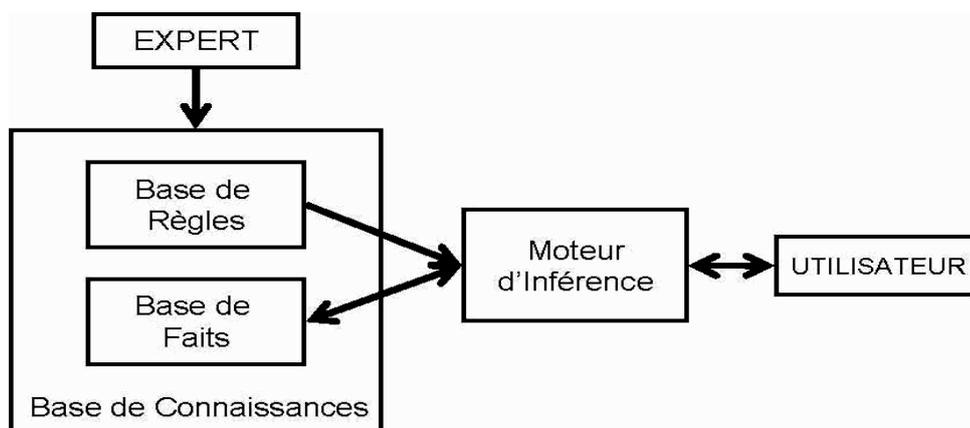


Fig1.20- Architecture d'un système expert.

Les principaux avantages des systèmes experts pour le diagnostic sont leur capacité à raisonner sous incertitude et leur capacité à apporter des explications des solutions fournies. La difficulté spécifique de leur mise en oeuvre est la formalisation de la démarche cognitive qui a pour objectif, à partir d'une situation donnée, de définir et de décrire le raisonnement associé.

2- La logique floue :

Les ensembles flous permettent de représenter et manipuler des données imprécises et incertaines. Leurs applications s'adressent à une aide à la décision [BAGHL99].

Un système flou réalise trois étapes essentielles :

- Fuzzification : On associe à chaque variable d'entrée un degré d'appartenance à un état déterminé à l'aide des fonctions d'appartenances définies dans le système.
- Inférence : les degrés d'appartenance de chaque variable à chaque état étant déterminé, on applique alors les règles d'inférence préalablement définies pour obtenir le degré d'appartenance des variables de sortie à chaque état.
- Défuzzification : À partir des règles d'appartenance des sorties, on applique un critère de décision pour calculer la sortie du système.

3- Séparateurs à vaste marge:

Les SVMs (Support Vector Machines ou Séparateurs à Vaste Marge) sont une méthode récente de classification supervisée, ils furent introduits par Vapnik en 1995[VAPN95]. Ils ont pour objectif de rechercher le meilleur hyperplan de séparation des données en deux classes (Figure 1.21). Cette méthode est basée sur l'utilisation de fonctions dites noyau (kernel) qui permettent une séparation optimale des données.

Les SVM ont montré leur efficacité dans de nombreux domaines d'applications comme la reconnaissance des formes [SHIGE05].

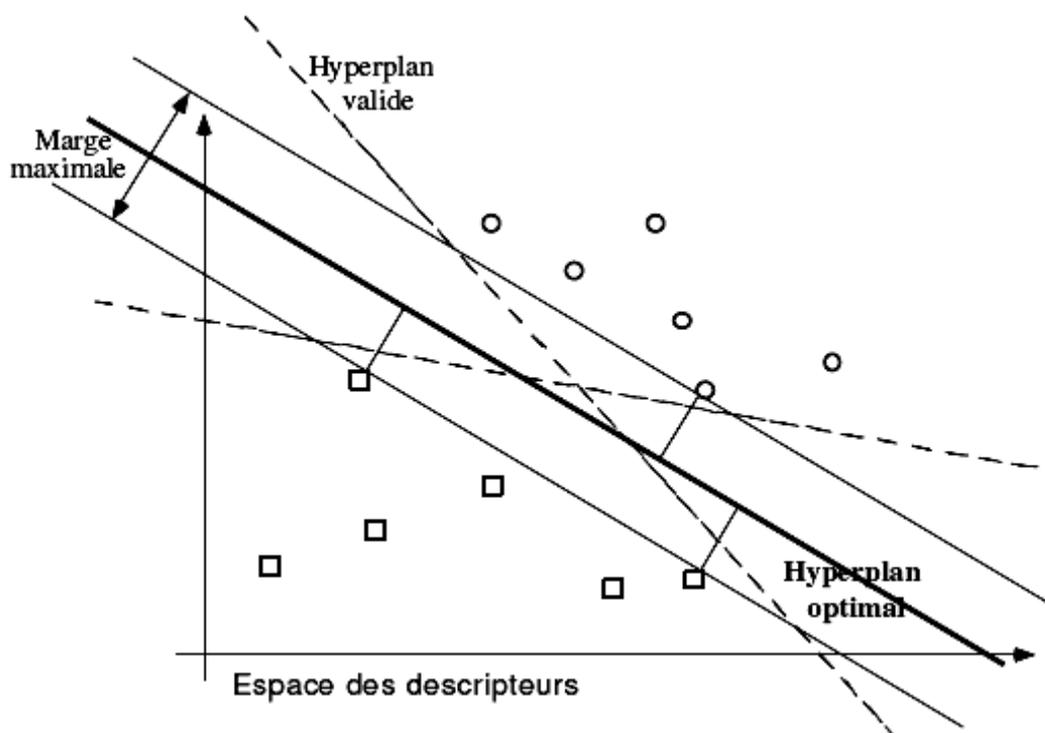


Fig1.21– Séparateurs à vaste marge.

4- Analyse discriminante:

L'analyse discriminante est une technique statistique qui vise à décrire, expliquer et prédire l'appartenance à des groupes prédéfinis (classes, modalités de la variable à prédire, ...) d'un ensemble d'observations (individus, exemples, ...) à partir d'une série de variables prédictives (descripteurs, variables exogènes, ...). On distingue généralement deux grandes approches:

- Approche prédictive: elle consiste à construire une fonction de classement (règle de décision) qui permet de prédire le groupe d'appartenance d'un individu à partir des valeurs prises par les variables prédictives. Les méthodes les plus utilisées en reconnaissance de formes sont le *Kppv* [MICH94].

- Approche descriptive ou "Analyse Factorielle Discriminante": Son objectif est de créer de nouvelles variables qui sont particulièrement efficaces pour séparer les classes. Ces nouvelles variables, dites "Axes discriminants", ou "Facteurs discriminants", sont obtenues comme combinaisons linéaires des variables originales. On peut citer plusieurs méthodes tel que le critère de Fisher [BENOU99].

5- Réseaux de neurones artificiels (RNA):

Les RNA sont des outils capables d'effectuer des opérations de classification. Ils sont composés d'éléments simples fonctionnant en parallèle. Ces éléments sont inspirés des systèmes nerveux biologiques. La fonction de chaque réseau est alors déterminée par les connections entre les éléments formant le réseau. Il est préférable, comme pour le cerveau humain, d'entraîner le réseau afin qu'il soit en mesure de réaliser une application particulière en ajustant les valeurs de ses connections. Le principal avantage des RNA par rapport aux autres méthodes est leur capacité d'apprentissage et de généralisation de leurs connaissances à des entrées inconnues.

Une des qualités de ce type d'outil est son adéquation pour la mise au point de systèmes de surveillance modernes, capables de s'adapter à un système complexe avec reconfigurations multiples. Les RNA peuvent être également implémentés en circuits électroniques, offrant ainsi la possibilité d'un traitement temps réel [RACO06].

Pour identifier des défauts dans un système, le diagnostic réalisé par réseaux de neurones doit disposer d'un nombre suffisant d'exemples de bon fonctionnement et de défauts pour pouvoir les apprendre. Pendant la phase d'apprentissage, les exemples sont présentés au réseau en entrée avec les étiquettes correspondantes à la sortie.

La figure 1.22 montre l'architecture générale d'une application de surveillance par reconnaissance des formes avec réseaux de neurones.

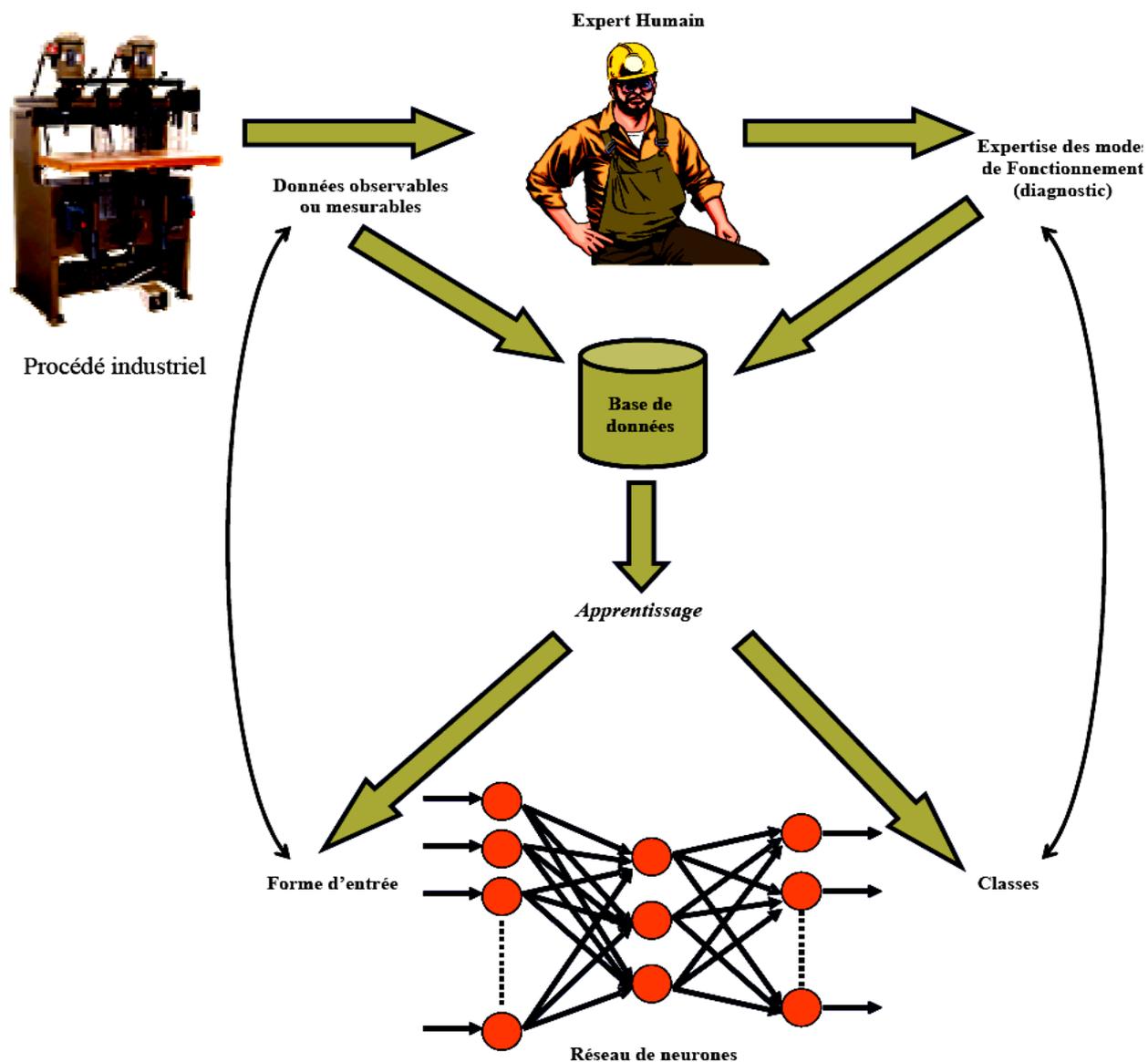


Fig1.22- Reconnaissance des formes par réseau de neurones [ZEMO03]

1.4-Conclusion:

L'objectif de ce chapitre était de donner un aperçu des différents défauts des machines tournantes et les techniques, d'analyse vibratoire, utilisées dans le diagnostic de ces défauts. Nous avons vu que l'apparition d'un défaut sur l'un des organes de la machine influe directement sur la forme du signal vibratoire reçu à partir de cette machine; chaque défaut se comporte différemment des autres, donc l'utilisation de plusieurs indicateurs est nécessaire pour faire un diagnostic efficace. Nous avons présenté, d'une part, les méthodes classiques qui se basent sur les techniques du

traitement de signal et, d'autre part, les méthodes automatisées qui se basent sur la classification automatique des signaux. Ces méthodes représentent d'excellents systèmes d'aide à la décision pour l'expert humain. Dans ce sens, le chapitre suivant sera consacré à l'état d'art de l'application de l'une des ces méthodes (réseaux de neurones) dans la surveillance des machines tournantes.

CHAPITRE2

ETAT D'ART DE L'APPLICATION

DES RESEAUX DE NEURONES

ARTIFICIELS (RNA) DANS

LE DIAGNOSTIC DES MACHINES

TOURNANTES

**CHAPITRE2: ETAT D'ART DE L'APPLICATION
DES RESEAUX DE NEURONES ARTIFICIELS (RNA)
DANS LE DIAGNOSTIC DES MACHINES
TOURNANTES**

2.1- Introduction :

Les réseaux de neurones artificiels (RNA) consistent en un ensemble d'outils et de méthodes de calcul. Ils sont caractérisés par leur capacité d'apprentissage, de généralisation, de mémorisation, de classification, d'adaptation et de décision. Ils sont appliqués dans divers domaines, tel que la reconnaissance de formes, le contrôle et la robotique. En maintenance industrielle, les réseaux de neurones sont utilisés pour résoudre les problèmes du diagnostic par la classification automatique des signaux et des formes correspondants aux différents états de fonctionnements normaux et anormaux des machines.

Dans la première partie de ce chapitre, nous présentons quelques notions concernant les types et les caractéristiques des réseaux de neurones artificiels. La deuxième partie contient une synthèse, de quelques travaux sur l'application des réseaux de neurones dans le diagnostic des machines tournantes.

2.2- Généralités sur les réseaux de neurones: [HAYK01]

2.2.1- Neurone biologique :

Dans le cerveau humain, le neurone est l'élément de base. Il reçoit des signaux en provenance de neurones voisins, les traite, engendre, conduit et transmet l'influx nerveux à d'autres neurones. La figure 2.1 représente les éléments d'un neurone biologique.

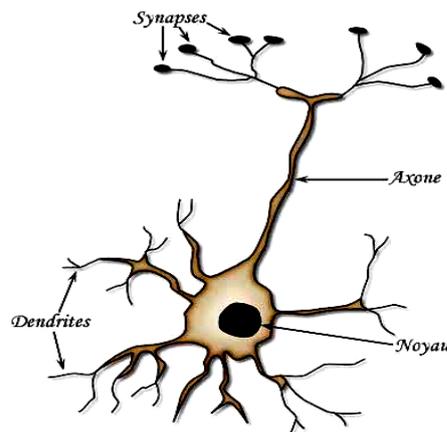


Fig2.1- Le neurone biologique

Le neurone est constitué de :

- **Les dendrites**: ce sont les récepteurs principaux du neurone pour capter les signaux qui lui parviennent.
- **Le corps cellulaire** : Il contient le noyau. C'est un sommateur à seuil. Il effectue une sommation des influx nerveux transmis par ses dendrites. Si la somme est supérieure au seuil, le neurone répond par un influx nerveux ou potentiel d'action qui se propage le long de son axone. Si la somme est inférieure au seuil, il reste inactif.
- **L'axone** : Il sert de moyen de transport pour les signaux émis par le neurone. Il se ramifie à son extrémité, là où il communique avec d'autres neurones.
- **Les synapses**: Ils permettent aux cellules nerveuses de communiquer entre elles. Les synapses se rencontrent surtout entre les axones et les dendrites.

2.2.2- Réseau formel :

Les réseaux de neurones formels sont à l'origine d'une tentative de modélisation mathématique du cerveau humain. Les premiers travaux datent de 1943 et sont l'oeuvre de MM. *Mac Culloch et Pitts*. Ils présentent un modèle assez simple pour les neurones et explorent les possibilités de ce modèle.

Le neurone représenté par la figure 2.2 se compose d'une cellule possédant plusieurs entrées et une sortie.

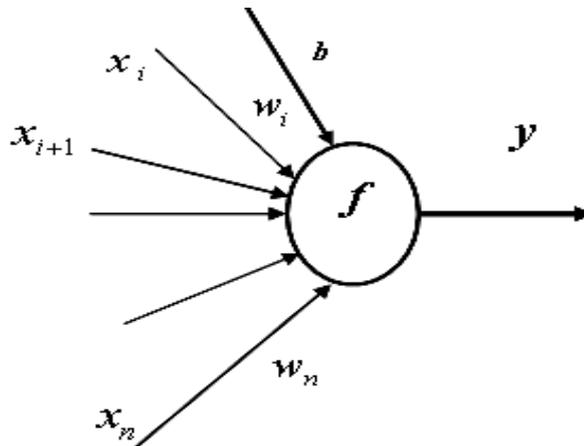


Fig2.2- Schéma d'un neurone formel

Avec :

$$y = f\left(\sum_{i=1}^n w_i x_i + b\right) \tag{2.1}$$

y : La sortie du neurone.

f : La fonction d'activation (ou de transfert).

x : Entrée du neurone.

w : Poids synaptique du neurone.

b : Biais.

Il existe de nombreuses formes possibles pour la fonction de transfert. Les plus courantes sont présentées sur la figure 2.3, d'autres sont présentées sur l'annexe 2. On remarquera qu'à la différence des neurones biologiques dont l'état est binaire, la plupart des fonctions de transfert sont continues, offrant une infinité de valeurs possibles comprises dans l'intervalle $[0, +1]$ ou $[-1, +1]$.

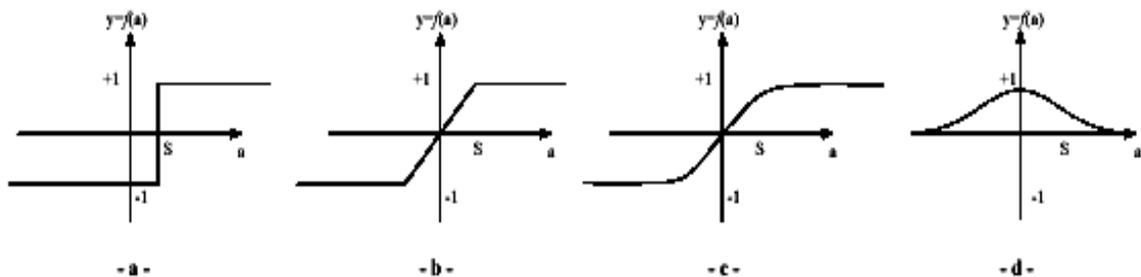


Fig2.3- Différents types de fonctions de transfert pour le neurone artificiel.

a) Fonction à seuil du neurone de Mc Culloch et W. Pitts (1949), **b)** Linéaire du modèle Adaline de Widrow et Hoff (1960), **c)** Sigmoïde d'un réseau perceptron multicouche de Rosenblatt (1962), **d)** Gaussienne du réseau de Moody et Darken (1989).

2.2.3- Apprentissage des Réseaux de neurones artificiels :

L'apprentissage des RNA, est une phase qui permet de déterminer ou de modifier les paramètres du réseau, afin d'adapter un comportement désiré. Plusieurs algorithmes d'apprentissage ont été développés depuis la première règle d'apprentissage de Hebb en 1949.

Les RNA se divisent en deux principales classes, les réseaux à apprentissage supervisés et les réseaux à apprentissage non supervisés.

- Pour les réseaux à apprentissage supervisés (Perceptron, Adaline,...etc.), on présente au réseau des entrées, et au même temps les sorties que l'on désirerait pour cette entrée. Le réseau doit alors se reconfigurer, c'est-à-dire calculer ses poids afin que la sortie qu'il donne corresponde bien à la sortie désirée.
- Pour les réseaux à apprentissage non supervisé (Hopfield, Kohonen,...etc.), on présente une entrée au réseau et on le laisse évoluer librement jusqu'à ce qu'il se stabilise.

L'apprentissage des RNA se fait selon la stratégie suivante:

- Etape 1:

- Les entrées d'un exemple sont injectées dans les cellules d'entrée du RNA.
- Le RNA calcule alors une sortie en fonction des états internes des neurones et des poids synaptiques des connexions.

- Etape 2:

- La sortie obtenue est comparée avec la sortie désirée pour l'exemple.
- On obtient alors une erreur, particulièrement importante en début d'apprentissage car la réponse du RNA est à ce stade purement aléatoire.

- Etape 3:

- Les poids synaptiques du RNA sont ensuite modifiés afin de réduire l'erreur calculée.

- Etape 4:

- On réitère cette opération un grand nombre de fois, et ce pour chaque exemple de la base, jusqu'à ce que le RNA converge vers une configuration qui lui permette de résoudre le problème à traiter.

2. 2.4- Architecture des RNAs:

Il existe différentes topologies de réseaux correspondant aux divers schémas géométriques de connexion entre les neurones. Les connexions entre les neurones qui composent le réseau décrivent la topologie du modèle. La figure 2.4 présente une taxonomie possible en terme d'architecture de réseaux. La différence majeure porte sur la possibilité d'avoir des boucles dans le réseau. Par ailleurs, on notera la possibilité d'avoir des «couches» de cellules.

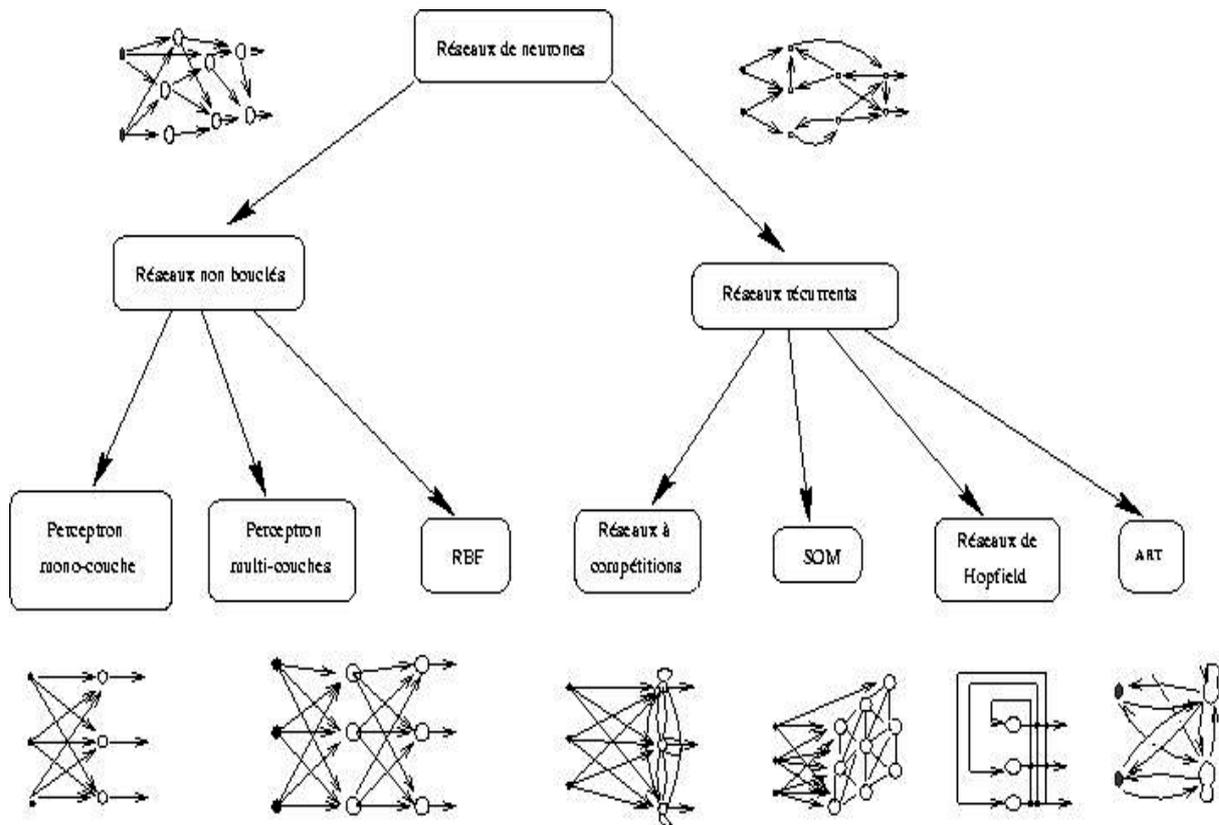


Fig.2.4- Différentes architectures des RNA.

RBF: Radial Basic Function.

ART: Adaptatif Resonance Theory.

SOM: Self Organized Map.

2.3- Bibliographie sur l'application des RNAs dans le diagnostic des machines tournantes :

Les réseaux de neurones par leurs caractéristiques- décrites précédemment- peuvent fournir une solution intéressante pour les problèmes de la maintenance des machines tournantes. De nombreux auteurs ont abordé ce sujet, on peut citer les travaux suivants:

ROEHL et al. [ROHE 90] essayent dans leur recherche, de développer une nouvelle approche dans la détection des défauts dans les machines tournantes. Cette approche est basée sur la combinaison des réseaux de neurones avec la logique floue. Ils ont appliqué cette méthode pour détecter les défauts de désalignement et du balourd dans un hydro- générateur. Les signaux vibratoires sont collectés à partir d'une paire de capteurs (horizontal et vertical) montés sur les paliers supportant l'arbre de la machine.

Les données du capteur horizontal sont combinées avec celles du capteur vertical pour former des orbites. Les orbites des vibrations construisent le vecteur d'entrée du réseau de neurone utilisé qui est de type ART floue (adaptive resonance theory). L'entrée du réseau contient 150 composantes. Le codage des neurones (étiquettes) de sortie est contrôlé par le paramètre (ρ) qui varie entre 0 et 1. Ils trouvèrent, que pour $\rho < 0.65$, le réseau ne distingue pas entre les deux fonctionnements (normal et anormal), et pour $\rho > 0.65$, il les classe correctement.

TIMO SORSA et al. [SORS91] utilisent les réseaux de neurones artificiels pour le diagnostic d'un réacteur thermique. Dans leur étude, ils utilisent le modèle mathématique qui simule le processus du réacteur. Un ensemble de 14 indicateurs correspondants aux dix défauts sont tirés à partir du modèle et leurs valeurs sont bruitées avec un bruit qui varie de 0% à 10%. La simulation est faite sur un ordinateur IBM PS/2 Modèle 70.

Ils normalisent les indicateurs entre -1 et 1 pour rendre la classification plus facile, et réduisent leur nombre par l'application de l'analyse en composante principale (ACP). Ils utilisent trois types de réseaux de neurones : Le perceptron avec une seule couche, le perceptron multicouches avec rétropropagation du gradient et le réseau counterpropagation, ce dernier est une combinaison entre le réseau de Kohonen et de Grossberg. Chaque réseau contient dans le vecteur d'entrée 14 neurones qui correspondent aux indicateurs, et dix neurones dans le vecteur de sortie qui correspondent aux différents défauts. La fonction d'activation des neurones de sortie est de type sigmoïde. Les données d'apprentissage contiennent 440 mesures, et les sorties désirées des réseaux sont fixées à 0 pour le fonctionnement normal et à 1 dans le cas de défauts. Pour le test des réseaux, si les valeurs des sorties sont supérieures à 0.5, alors il y a un défaut, dans le cas contraire, le fonctionnement est normal. Tous les réseaux sont entraînés jusqu'à un nombre d'itérations égale à 5000.

Les auteurs ont trouvé que :

- Le réseau perceptron avec une seule couche peut classifier sans difficultés plusieurs situations, mais quelques observations à l'entrée donnent des sorties qui sont très loin des valeurs désirées (étiquettes).
- Le réseau perceptron multicouches qui a la configuration suivante: au moins quatre neurones dans sa couche cachée, l'algorithme de l'apprentissage est la rétropropagation du gradient avec un coefficient d'apprentissage $\eta=0.95$ et un moment $\alpha=0,60$. Ce réseau permet une

- meilleure classification que celle du réseau précédent. La variation des paramètres du réseau n'augmente pas la vitesse de convergence de l'apprentissage, mais ceci n'est possible que par l'utilisation de la fonction tangente hyperbolique dans la couche cachée. l'augmentation du nombre des couches cachées rend la simulation plus lente.
- Le réseau counterpropagation est plus rapide en apprentissage, mais ne peut faire l'apprentissage de certains cas, donc il est inacceptable pour le diagnostic.
- Le réseau qui donne les meilleurs résultats est le réseau perceptron multicouche qui a la structure suivante **14-4-10** avec une fonction d'activation de type tangente hyperbolique dans sa couche cachée.

MATHEW J.BEOK et al. [BEOK93] utilisent les réseaux de neurones pour automatiser le diagnostic des machines. Le dispositif expérimental utilisé est un ventilateur qui permet d'étudier les défauts de déséquilibre de l'arbre et de la fissuration des ailettes. Ils placent deux accéléromètres, l'un horizontal et l'autre vertical sur l'arbre. Le ventilateur tourne à une vitesse constante. Les signaux prélevés par les accéléromètres sont amplifiés, puis passent par un analyseur FFT.

Dans la phase de prétraitement, ils extraient, à partir des spectres des signaux, les pics les plus significatifs. Le vecteur d'entrée du réseau de neurones contient 18 valeurs. Ces valeurs sont normalisées entre 0 et 1. Le réseau de neurones utilisé est un perceptron multicouche avec une seule couche cachée, qui contient un nombre de neurones varié entre 5 et 20. L'apprentissage du réseau est effectué par l'algorithme de la rétropropagation du gradient avec un coefficient d'apprentissage de 0.02 et un moment de 0.9. Les performances du réseau sont testées pour cinq états (différents défauts) du ventilateur.

En conclusion, tous les réseaux utilisés peuvent classifier avec succès les défauts et surtout le défaut du déséquilibre. Mais, pour augmenter les performances des réseaux de neurones, un filtrage des signaux est nécessaire.

A.C.MacCROMICK ET A.K.NANDI, [MACCR96a] appliquent les réseaux de neurones artificiels pour la classification automatique des états d'une machine, en combinant les différents moments d'ordre supérieur du signal temporel (vibrations).

Les données expérimentales utilisées pour ce travail sont obtenues à partir de l'installation expérimentale suivante :

Le banc d'essai peut être divisé en trois sous-ensembles : l'ensemble du dispositif, les capteurs et l'ordinateur. L'ensemble de machine utilisé se compose d'un moteur qui fait tourner un arbre avec un volant. Le volant a huit trous filetés dans lesquels, on peut poser des petits poids pour faire le déséquilibre de l'arbre. Un bloc de roulements avec un trou fileté, est placé au-dessus de l'arbre. Par ce trou une tige en laiton peut être placée, pour créer le défaut de frottement. L'arbre est tenu, en place, avec un bloc de roulement auquel des capteurs horizontaux et verticaux ont été attachés.

Les capteurs utilisés sont deux accéléromètres. Les signaux de sortie sont échantillonnés à 12kHz. Les données sont alors filtrées à une bande de largeur de 1.3kHz.

L'utilisation de ce dispositif donne la possibilité de créer quatre différents états de la machine :

1. **N-N**: Pas de défauts.
2. **N-R**: Seulement le défaut du frottement est appliqué.
3. **W-N**: Seulement le défaut de déséquilibre est appliqué.
4. **W-R**: Les deux défauts sont appliqués.

Ces conditions sont créées pour une gamme de différentes vitesses de machine variants de

77HZ aux *100HZ*. des vibrations.

Le réseau de neurones utilisé est un réseau perceptron multicouche avec un apprentissage de type rétropropagation du gradient. Ce réseau a quatre neurones dans la couche de sortie. Le nombre des couches cachées varie de 1 à 2, et le nombre des neurones qu'elles contiennent varie entre 6 et 11. Le réseau est entraîné pour un nombre d'itérations supérieur à 1000.

Les auteurs combinent les deux premiers moments de la série temporelle complexe avec les trois premiers moments de sa dérivée pour former le vecteur d'entrée du réseau. Les valeurs du vecteur d'entrée doivent être normalisées entre -1 et +1. Chaque série temporelle produit un ensemble de 24 observations, les huit premières observations utilisés pour l'apprentissage du réseau, et les seize restantes pour le test.

Ils ont trouvé que :

- Après 5000 itérations, le taux de succès de l'apprentissage est 100%, mais le taux de succès du test n'est que de 80.8%.
- Le premier moment de chaque série temporelle a une plus grande influence sur les résultats plus importante que les autres moments.

- Le réseau qui a deux couches cachées, la première contient 6 neurones et la deuxième contient 4 neurones, donne les meilleurs résultats car après 1000 itérations, il classe les données avec un succès de 100% pour l'apprentissage, et un taux de succès de 91.4 % pour le test.
- La combinaison des deux premiers moments de la série temporelle et les trois premiers moments de sa dérivée, avec les deux premiers moments de son intégrale, en utilisant un réseau de neurones avec deux couches cachées ayant six neurones dans chacune d'elles, donne un succès de 100 % pour l'apprentissage et de 99.4 % pour le test.

A.C.MacCROMICK ET A.K.NANDI [MACCR96b], dans cette étude, font une comparaison entre les RNA, et d'autres méthodes statistiques de la classification des défauts dans les machines tournantes. Ils utilisent les données décrites dans [MACCR96a]. Dans la phase du prétraitement, ils combinent les signaux mesurés par les deux capteurs pour former des séries temporelles complexes. Aussi, ils extraient à partir de ses signaux, la moyenne de l'amplitude pour le défaut du déséquilibre et la variance pour le défaut du frottement, ainsi que les moments.

Dans la phase du traitement (classification), ils utilisent deux types des réseaux de neurones: Le perceptron multicouches (MLP) avec dix neurones dans la couche d'entrée et quatre dans la couche de sortie. Ce réseau est entraîné jusqu'à 10000 itérations pour atteindre une erreur de 10^{-6} . Ils varient le nombre des couches cachées de 1 à 2, ainsi que leur nombre de neurones variant de 2 à 23 neurones. L'autre réseau est la fonction radiale de base (RBF) avec dix entrées et quatre sorties et un nombre de neurones variant entre 5 et 23. Les méthodes statistiques utilisées sont : L'analyse discriminante linéaire, le plus proche centroïde, les vecteurs canoniques et les plus proches voisins.

Ils trouvent que:

- Le temps d'apprentissage du réseau RBF est plus court que celui du réseau perceptron multicouche; mais, le réseau RBF nécessite un nombre de neurones plus grand que celui du réseau MLP pour classer les défauts avec succès.
- La méthode du plus proche voisin, est la seule, parmi les méthodes statistiques utilisées, qui a pu classer toutes les données avec succès; mais, elle prend un temps très grand par rapport à celui des deux réseaux de neurones MLP et RBF.

B.A.PAYA et al. [PAYA97] emploient les réseaux de neurones artificiels dans le diagnostic des machines tournantes, en utilisant la transformée en ondelettes, pour le prétraitement des données. Les signaux utilisés ont été prélevés sur un banc d'essai de défauts de roulements et d'engrenages. L'accéléromètre est placé sur le palier des roulements. Six états de fonctionnements ont été étudiés:

1. Roulement sans défauts, pignon sans défauts.
2. Roulement sans défauts, pignon avec défaut (a).
3. Roulement sans défauts, pignon avec défaut (b).
4. Roulement avec défauts, pignon sans défauts.
5. Roulement avec défauts, pignon avec défaut (a).
6. Roulement avec défauts, pignon avec défaut (b).

Le défaut (a) du pignon a été créé en soudant une petite masse de diamètre 2mm sur l'une des dents. Le défaut (b) du pignon a été créé par la fracture d'une dent.

Le défaut de roulement a été créé par une fracture de 1mm sur la bague intérieure du roulement. Pour chaque état, 30 signaux sont enregistrés, et chaque signal vibratoire contient 1024 points. Ces signaux ont été transférés sur un ordinateur, où ils sont transformés du domaine temporel au domaine des ondelettes en utilisant la transformée D4 (Daubechies d'ordre 4). Les dix amplitudes de la transformée en ondelettes les plus élevées, et les dix nombres d'ondelettes qui leurs correspondent sont choisis, pour former le vecteur d'entrée du réseau de neurones.

Le réseau de neurones utilisé est un réseau multicouche avec 20 neurones dans la couche d'entrée, 14 neurones dans la couche cachée et 6 neurones dans la couche de sortie. L'apprentissage du réseau est effectué par l'algorithme de rétropropagation du gradient. L'erreur quadratique est calculée à chaque 500 itérations. 48 signaux sont utilisés pour faire l'apprentissage du réseau et le même nombre est utilisé pour faire le test.

Les résultats obtenus ont montré que le réseau employé a classé tous les cas avec un taux de succès de 96%. La transformée en ondelettes, en la comparant avec d'autres méthodes comme FFT, est un outil puissant, pour extraire les informations utiles à partir des signaux vibratoires.

C. KIRKHAM et T. HARRIS [KIRKH97] essaient de développer et d'exécuter un système hybride de réseaux de neurones, utilisé comme outil de diagnostic pour l'identification des défauts des machines tournantes, en utilisant l'analyse des vibrations. Le système est un processus à deux étapes : la première étape détermine si

une mesure faite donne un état normal , en le comparant à un réseau de Kohonen qui a été entraîné sur les bonnes données d'état de fonctionnement de la machine.

Le réseau de la deuxième étape (type RBF) utilise les sorties de la première étape (les déviations de l'état normal) et classe les états de la machine (défauts).

Une fois la base de données est enregistrée sur un microordinateur, les valeurs nécessaires des fréquences fondamentales sont calculées. La vitesse et les amplitudes correspondantes forment les entrées aux réseaux de Kohonen.

Le système prototype est développé dans un logiciel, avec la considération de l'implémenter sur un matériel. Ce système a été utilisé dans un moulin à papier, ils ont trouvé les résultats suivants:

- *Étape 1* : Sensibilité = 83% - le nombre correctement prédit en tant que bon ou normal: Négatifs faux = 3%, Positifs faux = 14%.
- *Étape 2* : Spécificité = 100% - des nouveautés correctes, le total diagnostiqué dans la classe correcte de défaut.
- La base de connaissance est développée avec les données limitées disponibles. Les résultats sont basés sur l'exécution du système comparé au diagnostic d'un technicien sur les mêmes données. Aucune donnée des machines de moulin n'a été employée pour développer l'étape de base de connaissance du système. Par conséquent, tous les résultats sont pour des données non apprises de défaut.

Ce travail et ses résultats indiquent qu'il est possible qu'un système de classification puisse diagnostiquer des données précédemment non apprises des défauts des machines.

LR. PADOVSE et al [PAD99] appliquent les réseaux de neurones pour la détection des défauts dans les boîtes de vitesses. Le banc d'essai utilisé contient un moteur à courant alternatif (AC) et deux réducteurs de vitesses avec deux étages. Ils ont créé trois défauts sur les dents du pignon. En variant la vitesse entre 400 *rpm* et 1400 *rpm* avec un pas de 200, ils placent un accéléromètre vertical sur le palier pour mesurer les vibrations, et un microphone devant le front du pignon pour mesurer le bruit acoustique.

Dans la phase de prétraitement, ils font un filtrage passe –bas avec une fréquence de coupure

de 2KHZ et un échantillonnage avec une fréquence d'échantillonnage $f_e = 5,12KHZ$.

Leur base de données contient 324 signaux.

Dans la phase du traitement, ils utilisent un réseau de neurones est de type perceptron multicouches, avec deux différents algorithmes d'apprentissage. L'un est l'algorithme de Marquardt, et l'autre est l'algorithme du gradient conjugué. Ils utilisent 80% des signaux pour l'apprentissage et 20% pour le test.

La classification des signaux se fait selon deux stratégies. L'une est générale, un seul réseau classe les défauts. L'autre est en cascade: à deux réseaux, le premier classe les vitesses, et le deuxième classe les défauts. Ils utilisent deux vecteurs d'entrées pour les réseaux de neurones. Un vecteur contient les valeurs du kurtosis et du RMS, et l'autre contient les pics du spectre (Welch).

Ils ont trouvé, que pour un nombre d'itérations égale à 500, et une erreur de 0.01:

- L'entrée qui contient les valeurs du spectre donne des résultats meilleurs que celle du vecteur des statistiques.
- L'algorithme de l'apprentissage conjugué du gradient est meilleur que l'algorithme de Marquardt.
- Le réseau en cascade a des meilleurs performances que le réseau général.

D.YANG et J. PENMAN [YANG00] utilisent les données du courant électriques du moteur et celles des vibrations pour le diagnostic des roulements par la méthode des réseaux de neurones artificiels. Leur dispositif expérimental contient un moteur à induction 15KW/1500rpm entraînent un alternateur 11KVA par l'intermédiaire d'un accouplement flexible. Ils ont étudié quatre états des roulements:

- N-N: roulement sans défauts et sans charge.
- F-N: roulement avec défaut et sans charge.
- N-F: roulement sans défauts et avec charge.
- F-F: roulement avec défaut et avec charge.

Un accéléromètre est bien placé avec un goujon sur la surface supérieure du roulement d'essai pour prendre les signaux de vibration. Un capteur de courant est placé sur l'une des lignes des câbles pour collecter les signaux du courant de moteur. Tous les signaux mesurés sont échantillonnés à 12.5 KHZ avec un nombre de points de 8192.

Dans la phase du prétraitement, ils ont d'abord converti les signaux échantillonnés du domaine temporel au domaine fréquentiel par la méthode de Welch. Les 8192 points considérés sont divisés en 2 fenêtres coulissantes avec un chevauchement de 75 %. Le signal original est modifié par une fenêtre de Hanning. Puis, ils extraient les 18

fréquences les plus importantes pour former le vecteur d'entrée du réseau de neurones. Ces valeurs doivent être normalisées. Un total de 160 ensembles de formes de données de courant et de vibration sont rassemblés. La moitié des données est employée pour l'apprentissage des réseaux de neurones, et l'autre moitié est employée pour tester les performances du réseau.

Dans la phase du traitement, ils ont utilisé un réseau multicouche avec une seule couche cachée pour les données des vibrations et la même chose pour les données du courant. Ces réseaux contiennent 18 neurones dans la couche d'entrée, 4 neurones dans la couche de sortie et différents neurones dans la couche cachée. Ils ont utilisé plusieurs architectures pour une erreur égale à 10^{-10} , et un nombre d'itérations de 10000. Les coefficients d'apprentissage et de moment ont été pris respectivement 0.01 et 0.95. Ils ont trouvé que:

- Dans la phase de l'apprentissage, tous les réseaux donnent un taux de succès égal 100%.
- Dans la phase du test, les réseaux sont utilisés pour différencier entre les différents états du roulement. Les résultats trouvés, montrent que les réseaux, contenant les données des vibrations, ont pu distinguer entre les quatre états avec un taux de succès de 100 %. La même constatation est faite pour les réseaux contenant les données du courant (taux de succès de 100 %), sauf pour l'état F-F, le taux de succès de 97.5%.
- La comparaison entre les performances des réseaux montre que les réseaux basés sur les données de vibrations sont plus performants que les réseaux basés sur les données du courant; mais la surveillance par le courant fournit une méthode alternative pour diagnostiquer l'état de roulement qui est plus simple à appliquer.

L.B.JACK, A.K.NANDI [JACK00] utilisent les algorithmes génétiques pour sélectionner les données utilisées comme entrées des réseaux de neurones. Ils ont utilisé un banc d'essai qui était réalisé pour étudier les défauts de roulements. Ils étudient six états de roulements:

- Deux normaux ; un roulement est nouveau et l'autre est en état légèrement usé.
- Quatre anormaux (avec défauts):
 - Défaut sur la bague intérieure.
 - Défaut sur la bague extérieure.
 - Défaut sur la cage.

- Défaut sur les éléments roulants.

Chaque signal prélevé par l'accéléromètre est filtré par un filtre passe-bas avec une fréquence de coupure de 18,3 KHZ. Puis, il est échantillonné avec une fréquence de 48 KHZ. Cette opération est répétée pour seize différentes vitesses, ce qui donne pour les six états de roulements une base de données de 960 signaux.

L'entrée des réseaux de neurones est formée des moments et des cumulants d'ordre supérieur (90 caractéristiques), et 66 valeurs correspondantes aux pics des spectres. Le réseau de neurones utilisé est de type perceptron multicouches avec:

- Une seule couche cachée, en variant le nombre des neurones de 2 à 15.
- Fonction de transfert de type logistic.
- Nombre d'itérations maximal égale à 350.
- Le type de l'algorithme d'apprentissage est la rétropropagation du gradient.
- Six neurones dans la couche de sortie, correspondants aux six états de roulements étudiés.
- 40% des signaux sont utilisés pour l'apprentissage du réseau, et les 60% restants sont utilisés pour le test.

Les auteurs trouvent que:

- La performance du réseau sans algorithmes génétiques avec l'entrée qui contient les valeurs du spectres, est meilleurs que celles de l'entrée qui contient les moments et les cumulants.
- L'utilisation des algorithmes génétiques réduit le nombres des entrées de 156 à 6, ce qui permet de réduire le temps de calcul et donc d'augmenter la performance du réseau. Le réseau de neurones idéal à 9 neurones dans sa couche cachée.

J.Mahfoudi et al. [MAHF00] appliquent les réseaux de neurones pour la détection et l'identification des défauts combinés dans les boites de vitesses.

Le banc d'essai est constitué de deux axes montés sur deux roulements et couplés aux deux engrenages. L'axe d'entrée est entraîné par un moteur D.C 120 KW. L'axe de sortie est freiné par un moteur C.C.R. La vitesse de rotation varie entre 0 et 6000 *tr/mn*. Le couple résistant a été appliqué indépendamment à partir de 0 à 150 *Nm*. Les engrenages d'essai sont 36/38 dents avec un rapport de contact de 1.3 et un module de 2 *mm*. Les roulements d'essai sont de type *SKF 6012* pour le roulement à billes et de type *SKF NU1013* pour le roulement à rouleaux. Un défaut uniformément distribué sur l'engrenage a été fait par l'écaillage. Des défauts locaux sur les deux roulements à

billes et à rouleaux ont été créés par un enlèvement de métal, sur la bague intérieure, à l'aide d'une meuleuse.

Des accéléromètres sont placés sur les paliers près de chaque élément défectueux. Les signaux sont enregistrés à l'aide d'une carte d'acquisition et d'un micro-ordinateur après un filtrage de 20 KHZ. La fréquence d'échantillonnage est de 50 KHZ, cette fréquence tient compte de tous les phénomènes étudiés (défauts sur roulements et sur engrenage).

Les signaux mesurés sont alors traités dans le domaine fréquentiel. Un filtre simple avec la fréquence centrale réglable est utilisé pour l'analyse à bande étroite. Dans cette étude, un filtre avec une bande de largeur constante, d'un pourcentage (6%), est appliqué afin d'augmenter la visualisation. Les signaux sont également traités dans le domaine des cepstres, parce que les défauts locaux induisent des bandes latérales autour des fréquences d'engrènement de l'engrenage. Les spectres sont obtenus à partir d'une moyenne de 16 transformations successives de 1024 points. Les conditions de fonctionnement sont 3 vitesses (1000, 2300 et 4300 *tr/mn*) et 2 charges (60 et 120Nm). Plusieurs configurations sont considérées en relation avec les défauts étudiés (défaut distribué sur l'engrenage, défauts locaux sur les roulements et leurs combinaisons). Les indicateurs choisis sont :

- les 5 premières harmoniques de la fréquence d'engrènement et l'énergie correspondante dans le domaine des spectres.
- Les deux premières harmoniques du roulement à rouleaux et du roulement à billes dans le domaine des cepstres.

Trois réseaux sont employés pour l'automatisation de la détection des défauts, un réseau pour chaque élément défectueux. Les entrées sont les 14 indicateurs choisis pour chaque défaut présenté précédemment. La sortie est un neurone simple qui peut prendre les valeurs (0 ; 1) selon la présence (1) ou pas (0) du défaut considéré. Les couches cachées n'ont pas été considérées pour tester le processus mais en particulier pour calculer les poids (w_i). Huit modèles sont employés pour le procédé d'étude en utilisant l'apprentissage par la rétropropagation du gradient.

Les résultats obtenus montrent que tous les modèles utilisés ont pu classifier les différents états de fonctionnement avec succès. Aussi, l'optimisation des réseaux de neurones permet de déterminer les indicateurs nécessaires pour la détection des défauts combinés des roulements et des engrenages.

B.Samanta [SAMA04], dans son étude, compare entre les performances des réseaux de neurones artificiels (RNAs) et celles des séparateurs à vaste marge (SVM), utilisés pour la détection des défauts d'engrenages. Les signaux vibratoires sont mesurés à partir de sept accéléromètres montés sur une pompe entraînée par un moteur électrique, par l'intermédiaire d'un réducteur de vitesse à deux étages. Des mesures séparées ont été obtenues, pour deux machines identiques, l'une avec un défaut de piqures sur les deux paires d'engrenage, et l'autre sans défauts. Les signaux ont été traités par un filtre passe-bas, avec une fréquence de coupure de 1 et 5 KHZ, respectivement. Le nombre d'échantillons rassemblé pour chaque canal est 77824 points.

Le vecteur de forme utilisé comme entrée aux RNAs et aux SVMs contient 45 caractéristiques normalisées (indicateurs):

- **1-9:** sont les valeurs de la moyenne, la valeur efficace, la variance, skewness, kurtosis, et les moments centrés de l'ordre 5 à l'ordre 9.
- **10-27:** les valeurs calculées des dérivés et des intégrales de chaque signal.
- **28-45:** signaux filtrés par des filtres passe-haut et passe-bas avec une fréquence de coupure de 640 HZ, pour fréquences d'échantillonnage de 3.2 et 12.8 kSa/s.

La procédure de l'extraction des caractéristiques a été répétée pour deux conditions de charge (maximale et minimale), pour deux taux d'échantillonnage (haut et bas), et deux états d'engrenage (défectueux et normal). Chaque accéléromètre donne 38 signaux (20 pour l'apprentissage et 18 pour le test). Donc, il obtient un vecteur forme de $45 \times 266 \times 2 \times 2 \times 2$.

Le réseau de neurones utilisé est un perceptron multicouches avec la configuration suivante:

- Une seule couche cachée, dont le nombre des neurones varie entre 10 et 30.
- Le nombre des entrées varie entre 3 et 45, le nombre des neurones dans la couche de sortie varie entre 1 et 2.
- Les fonctions d'activation sont: tan- sigmoïde et logistic.
- Le réseau de neurone a été créé et implémenté en utilisant MATLAB neural networks toolbox avec l'algorithme de rétropropagation du gradient et un algorithme d'apprentissage de Levenberg–Marquardt.
- La fonction de performance est MSE (MSE= erreur = 10^{-6})
- Un gradient minimal de 10^{-10} , et un nombre maximal d'itérations égal 500.

Pour les SVMs, il a choisi le paramètre Kernel $C=100$, σ et les caractéristiques ont été sélectionnées par l'approche des algorithmes génétiques.

Samanta a trouvé que:

- La classification des SVMs sans algorithmes génétiques (AG), est meilleure que celle des RNAs.
- Pour une sélection basée sur les AG, les performances des deux classifieurs se rapprochent de 100%.
- Aussi, le temps d'apprentissage de SVMs est un peu plus petit que celui des RNAs.

2.4- Conclusion:

Ce chapitre a été consacré, à la présentation d'un état d'art de l'application des réseaux de neurones artificiels pour le diagnostic des machines tournantes. Après avoir introduit leur concept de base, nous avons présenté une synthèse bibliographique de quelques travaux sur l'application des réseaux de neurones pour le diagnostic et la détection des défauts dans les machines tournantes. Les réseaux de neurones ont alors une large utilisation dans la maintenance industrielle. Différents types de réseaux sont utilisés, et la classification des différents états des machines peut être réalisée soit à partir des caractéristiques des signaux réels prélevés des machines, soit à partir des signaux simulés générés par des modèles. Cependant, pour exploiter un système capable de faire un diagnostic précoce, il reste à bien choisir les données d'apprentissage et du test, ainsi que le type et les paramètres du réseau. Il existe différentes techniques pour optimiser la structure du réseau ainsi que le nombre de signaux d'apprentissage.

Le chapitre suivant décrit, l'utilisation de l'un de ces réseaux, appelé le perceptron multicouches pour le diagnostic des roulements et d'engrenage, par l'analyse vibratoire, ainsi que son implémentation sous Matlab. Plusieurs configurations du réseau et des bancs d'essais seront étudiées.

CHAPITRE 3

DEMARCHE EXPERIMENTALE

CHAPITRE 3: DEMARCHE EXPERIMENTALE**3.1 - Introduction :**

Nous avons vu dans le chapitre précédent, que les réseaux de neurones artificiels et surtout le réseau appelé perceptron multicouche (PMC) ont une large application dans la maintenance industrielle des machines tournantes. Dans ce chapitre, nous essayons d'utiliser ce réseau comme un outil de diagnostic des machines tournantes (roulements et engrenages). Nous étudions différentes configurations en variant les caractéristiques du réseau, et les indicateurs de surveillance formant le vecteur forme. Ces indicateurs sont extraits à partir des signaux réels de roulements et d'engrenages.

Ce chapitre est structuré en trois parties. La première partie contient un aperçu théorique sur le réseau perceptron multicouche, et son utilisation dans le diagnostic des machines tournantes.

La deuxième partie est consacrée à l'implémentation des réseaux de neurones utilisés comme un outil de diagnostic sous MATLAB, en précisant le choix des différents paramètres du réseau.

Dans la troisième partie, nous présentons les différents bancs d'essai et les signaux vibratoires prélevés.

3.2 - Réseau perceptron multicouches:

Le perceptron multicouches est le plus simple et le plus connu des réseaux de neurones. La structure, présentée par la figure 3.1, est relativement simple : une couche d'entrée, une couche de sortie et une ou plusieurs couches cachées. Chaque neurone n'est relié qu'aux neurones des couches précédentes, mais à tous les neurones de la couche suivante. La fonction de transfert utilisée est en générale une somme pondérée.

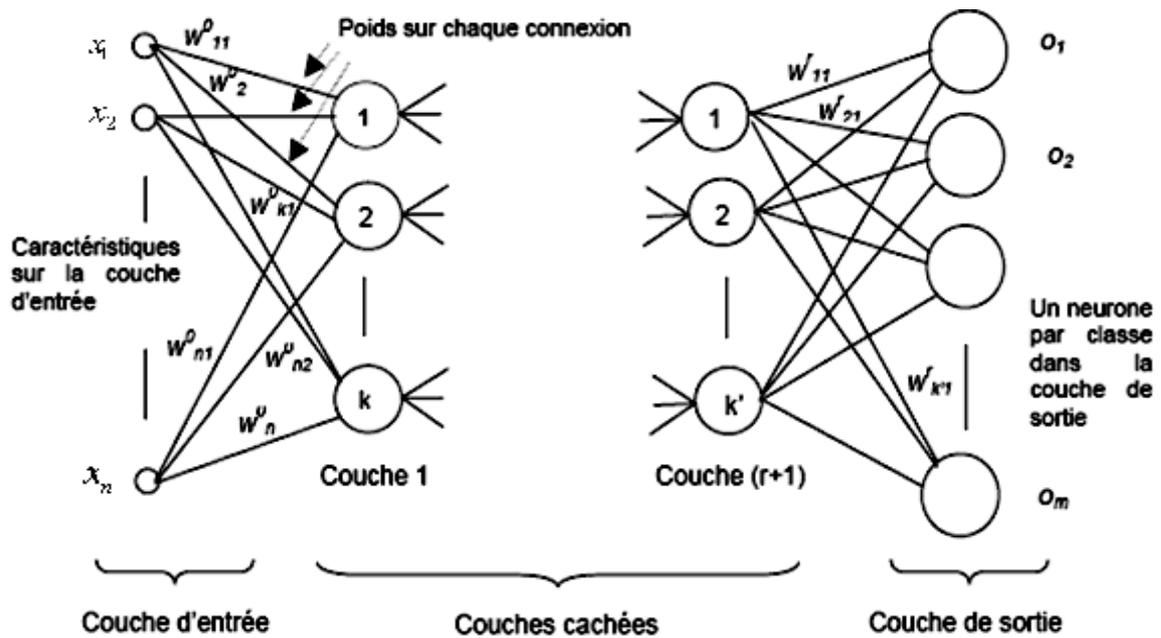


Fig. 3.1- structure d'un réseau perceptron multicouches.

- X : est le vecteur d'entrée du réseau.
- O : est le vecteur de sortie du réseau.
- W : poids synaptiques.

3.2.1 - Neurones d'entrée :

La première couche est appelée couche d'entrée. Elle recevra les données source que l'on veut utiliser pour l'analyse. Dans le cas de l'aide au diagnostic industriel, cette couche recevra les indicateurs de surveillance. Sa taille est donc directement déterminée par le nombre de variables d'entrées.

3.2.2 - Neurones cachés :

La seconde couche est une couche cachée, en ce sens qu'elle n'a qu'une utilité intrinsèque pour le réseau de neurones et n'a pas de contact direct avec l'extérieur. Les fonctions d'activation sont en général non linéaires sur cette couche mais il n'y a pas de règle à respecter. Le choix de sa taille n'est pas implicite et doit être ajustée. Souvent, il sera préférable pour obtenir de bons résultats, d'essayer le plus de tailles possibles.

3.2.3 - Neurones de sortie :

La troisième couche est appelée couche de sortie. Elle donne le résultat obtenu après compilation par le réseau de données d'entrée dans la première couche. Dans le

cas de l'aide au diagnostic industriel, cette couche donne l'état de la machine. Sa taille est directement déterminée par le nombre de classes (ou modes de fonctionnement).

3.2.4 - Apprentissage du réseau perceptron multicouches :

L'apprentissage du perceptron multicouches est supervisé, il consiste à adapter les poids des neurones de manière à ce que le réseau soit capable de réaliser la tâche demandée.

La méthode classique pour l'apprentissage du perceptron multicouche est l'algorithme de rétro-propagation, qui a été développé en particulier par Rumelhart et Parkenet le Cun en 1985. Cet algorithme repose sur la minimisation de l'erreur quadratique entre les sorties calculées et celles souhaitées. Le terme rétro-propagation du gradient provient du fait que l'erreur calculée en sortie est transmise en sens inverse vers l'entrée.

En résumé, nous utiliserons, pour l'apprentissage du réseau, l'algorithme de rétropropagation avec minimisation du gradient d'erreur qui est défini par les étapes suivantes [BISH95] :

1. Initialiser les poids à des petites valeurs et les seuils du réseau.
2. Insérer à l'entrée du réseau une observation (exemple) de la base de données en forme de vecteur de caractéristiques, puis calculer sa valeur d'activation et sa fonction d'activation :

$$net_i = \sum_j w_{ij} o_j \quad (3.1)$$

$$o_i = f(net_i) \quad (3.2)$$

Avec:

f : La fonction d'activation.

o_i : La sortie du neurone i .

3. Évaluer l'erreur des sorties du réseau :

$$\delta_j = (d_j - o_j) o_j (1 - o_j) \quad (3.3)$$

d_j : La valeur de l'erreur désirée.

4. Ajuster les poids:

$$\Delta w_{ij} = \eta \delta_j o_i \quad (3.4)$$

η : La valeur de taux d'apprentissage.

5. Évaluer le signal d'erreur pour chaque couche cachée :

$$\delta_j = o_j (1 - o_j) \sum_k w_{jk} \delta_k \quad (3.5)$$

6. Ajuster les poids de la couche cachée en utilisant la formule (3.4).

7. Répéter les étapes 2 à 6 pour l'ensemble des observations de la base d'apprentissage tant que le critère d'arrêt n'a pas été atteint.

La formule de calcul de l'erreur quadratique de l'ensemble de la base d'apprentissage définie comme suit

$$E_p = \frac{1}{2} \sum_p (\sum_k d_{pk} - o_{pk})^2 \quad (3.6)$$

Où p est l'indice d'un exemple de la base et k est l'indice du nœud de sortie. L'objectif est de minimiser cette erreur.

3.3 - Déroulement de la classification par le PMC :

Le traitement (classification) des différents états d'une machine par la méthode des réseaux de neurones artificiels se fait selon l'organigramme présenté par la figure 3.2 :

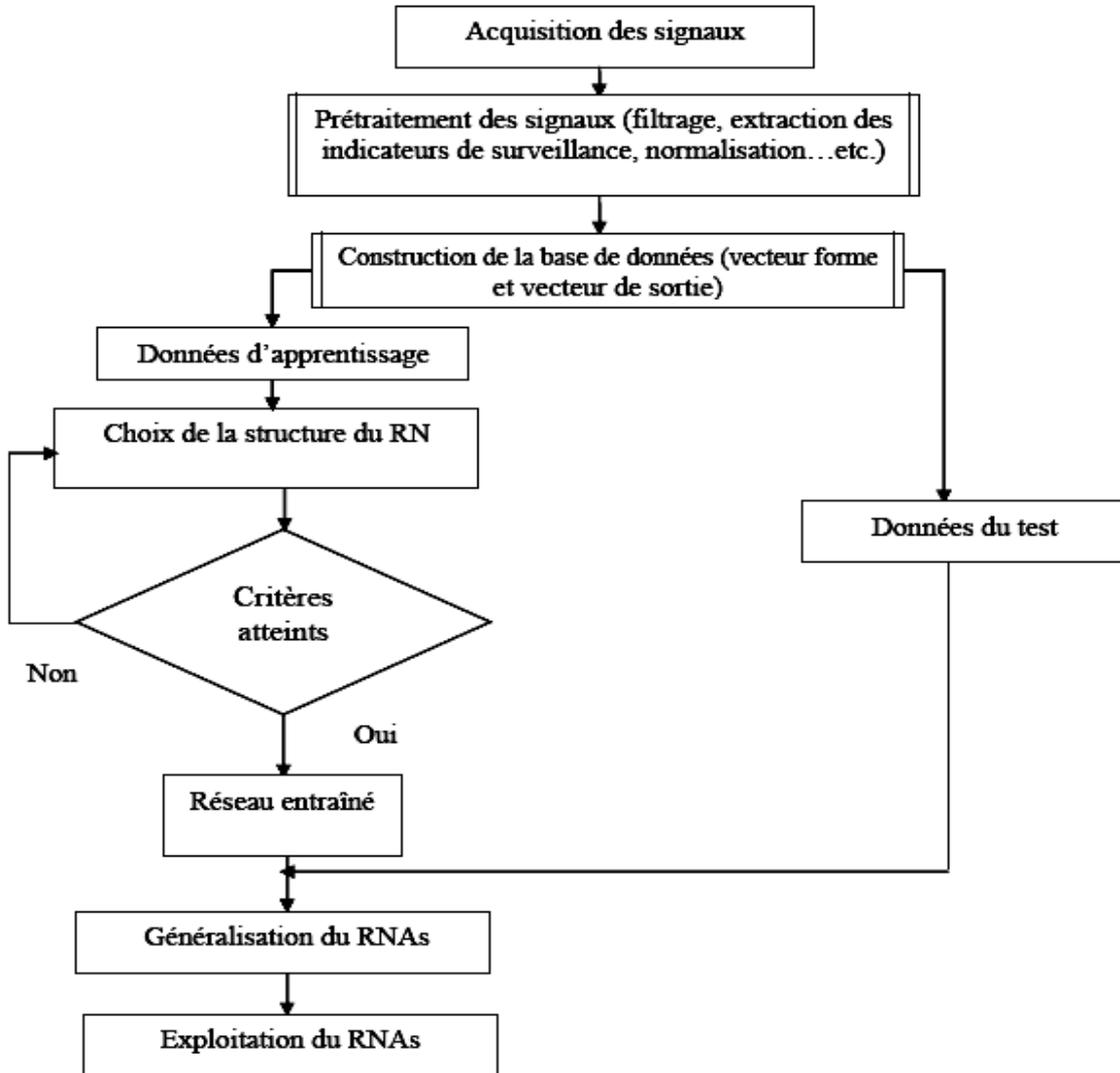


Fig.3.2 - Organigramme du déroulement de la classification par PMC.

3.3.1 - Acquisition des signaux :

Dans cette étape, on collecte (enregistrer) les signaux vibratoires correspondant aux différents états du dispositif étudié (la machine).

3.3.2 - Prétraitement des données :

Un prétraitement des données (entrées et sorties) est nécessaire pour éliminer toutes sorte d'informations inutiles existantes dans les différents signaux. Ainsi, cette étape consiste à faire un prétraitement sur la base de données (signaux), en utilisant les

différentes méthodes du traitement de signal comme le filtrage, l'analyse spectrale,...etc.

Aussi, on procède à une normalisation ou une sélection des données ; pour extraire les indicateurs les plus pertinents.

3.3.3 - Construction du vecteur forme :

Le prétraitement des données permet d'extraire les indicateurs de surveillance correspondant à l'état de la machine. Ces indicateurs forment les vecteurs (matrice) de la base d'apprentissage et du test du réseau.

3.3.4 - Apprentissage et test du réseau :

Une partie de la base des données est utilisée pour faire l'apprentissage du réseau, et l'autre est utilisée pour faire le test.

- L'apprentissage consiste à chercher une configuration optimale des paramètres du réseau (poids, biais,...etc.) correspondant à un taux de succès élevé.
- Une fois le réseau entraîné, il faut toujours procéder à des tests afin de vérifier que notre réseau réagit correctement. Il y a plusieurs méthodes pour effectuer une validation : validation directe, validation croisée...etc.

3.3.5 – Exploitation du réseau :

Après la détermination de la configuration optimale du réseau, ce dernier devient capable à classer d'autres états de la machine et décide l'existence ou non des défauts.

3.4 - Implémentation du PMC sous MATLAB :

L'implémentation du réseau PMC sous MATLAB se fait selon le schéma suivant :

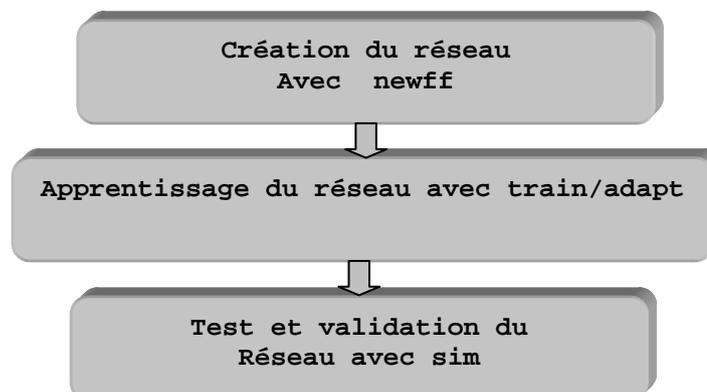


Fig3.3- Etapes pour implémenter un PMC sous MATLAB

3.4.1- Création du réseau :

La commande **newff** crée un réseau de neurones artificiel de type perceptron multicouche.

Syntaxe:

```
net= newff (PR, [S1, S2,... Si], {TF1 TF2 TFNL}, BTF, BLF, PF)
```

PR: matrice des valeurs minimales et des valeurs maximales du vecteur d'entrée P.

Si : nombre des neurones dans la $i^{\text{ème}}$ couche (couches cachées et couche de sortie).

TFNL: fonction de transfert dans la $i^{\text{ème}}$ couche. (Tansig, purelin,...etc.)

BTF: fonction d'apprentissage par rétro-propagation du réseau. (trainlm,... etc.)

BLF : type d'algorithme utilisé (learngdm,... etc.)

PF : fonction de performance (mse, mae...etc.)

3.4.2- Apprentissage du réseau :

L'apprentissage se fait avec la fonction **train** pour l'apprentissage de type batch, c'est-à-dire, le calcul de l'erreur se fait après la présentation de tous les exemples d'apprentissage. La syntaxe est :

```
net =train (net, P, T).
```

Avec : **net:** réseau de neurones créé.

P: matrice des observations (entrée du réseau).

T: matrices des étiquettes (sortie du réseau).

Ou par **adapt** pour l'apprentissage de type incrémental, c'est-à-dire, l'erreur est calculée après chaque exemple. La syntaxe est la suivante :

```
Net=adapt (net, P, T).
```

Avant d'exécuter les commandes **train** et **adapt**, il faut d'abord, entrer les paramètres d'apprentissage du réseau.

```
net.trainParam.lr = taux d'apprentissage.
```

```
net.trainParam.show = fréquence d'affichage.
```

```
net.trainParam.epochs = nombre max des itérations.
```

```
net.trainParam.goal = Erreur minimale.
```

3.4.3- Test et validation du réseau :

Le test du réseau se fait avec la commande **sim** qui a la syntaxe suivante :

$Y = \text{sim}(\text{net}, P)$

Y: sortie calculée (réelle).

3.5 - Choix des paramètres du réseau :

Le succès d'un réseau de neurones artificiel, utilisé pour résoudre un problème de classification, est en relation directe avec la configuration du réseau. Car une mauvaise configuration (choix des paramètres) rend l'apprentissage plus lent ou sature le réseau, ou encore crée le sur-apprentissage du réseau. Ce qui n'est pas souhaitable, et surtout dans les problèmes du diagnostic industriel qui nécessite une décision rapide et précise sur l'état de santé des machines.

3.5.1- Type du réseau :

La synthèse bibliographique effectuée à la section 2.3, nous permet de choisir le type du réseau qui est le perceptron multicouche, parce qu'il est capable de résoudre les problèmes non linéaires (le cas de plusieurs classes non linéairement séparable). Aussi, il s'implémente facilement sous MATLAB.

3.5.2 - Prétraitement des données :

Les données d'apprentissage ne sont pas présentées directement aux RNA, mais plutôt traitées d'avance. Le prétraitement des données (entrées/sorties) consiste à [KECM01] :

a- Normaliser les valeurs de ces données, parce que les vecteurs d'entrées et de sorties des réseaux de neurones artificiels contiennent généralement des valeurs différentes. Et s'il existe un grand décalage entre eux, le réseau risque de se saturer ou d'échouer dans la classification. Dans ce cas, une normalisation (rendre les valeurs très proches) des données est fortement demandée.

Il existe plusieurs méthodes pour normaliser un vecteur de forme, comme la méthode de normalisation par l'écart type et la moyenne.

b- Réduction de la taille des vecteurs d'entrées et de sorties, qui permet de construire des RNAs, avec moins de paramètres et mieux convient à apprentissage/généralisation. La réduction de la taille se fait par la méthode des algorithmes génétiques, analyse en composante principale (ACP)...etc.

3.5.3 - Nombre de couches cachées :

Une des premières décisions à faire, est de déterminer combien de couches cachées sont nécessaires afin d'avoir un bon modèle. D'abord, on devrait affirmer qu'il n'y a aucun besoin d'avoir plus de couches cachées. Cette réponse est soutenue par les résultats théoriques et par beaucoup de simulations dans différents champs de technologie [THEOD03].

3.5.4 - Nombre de neurones :

Bien que le nombre de neurones dans les couches d'entrée et de sortie d'un réseau de neurones artificiels, puisse facilement être décidé pour le problème étudié. Il n'y a aucun critère pour déterminer le nombre des neurones convenable des couches cachées. Avec un nombre limité de neurones, le réseau ne sera pas performant sur l'apprentissage tandis qu'avec un nombre trop important de neurones il sera difficile à entraîner. On juge la performance d'un réseau sur sa capacité à généraliser les résultats et non pas sur sa capacité à mémoriser les résultats. C'est pour cela qu'il n'est pas toujours profitable d'utiliser un réseau avec un nombre élevé de neurones [MARQ01] .

3.5.5 - Initialisation des poids:

Le choix des valeurs initiales des poids synaptiques pour l'apprentissage par l'algorithme de rétro-propagation du gradient, est l'un des aspects les plus importants qui affectent la procédure d'apprentissage des RNA. Pratiquement, il est nécessaire d'initialiser les poids aléatoirement avec des petites valeurs, par exemple, entre -0,5 et 0,5 [GUPT 03].

3.5.6 - Taux d'apprentissage :

Le taux d'apprentissage η qui détermine l'ordre des incréments de poids à chaque étape d'ajustement, affecte certainement les performances de l'apprentissage des RNAs. Généralement, on peut dire qu'une valeur grande du taux d'apprentissage peut causer des oscillations instables. Tandis qu'un taux d'apprentissage η plus petit peut ralentir la procédure d'apprentissage. Dans l'algorithme conventionnel de la rétro-propagation du gradient, on suppose que le taux d'apprentissage est une valeur fixe et doit être uniforme avec tous les poids [GUPT03].

3.5.7 - Fonction d'activation :

Comme les autres questions pratiques concernant les réseaux de neurones artificiels, il n'y a aucune réponse précise au sujet du choix de la fonction d'activation. Les fonctions d'activation les plus utilisées sont la fonction sigmoïde (pour les couches cachées) et la fonction linéaire (pour la couche de sortie). Toutes les deux fonctions satisfont le critère de base; elles sont différentiables [DUDA03].

3.5.8 - Algorithme d'apprentissage :

Il est très difficile de savoir quel algorithme d'apprentissage sera le plus rapide en convergence pour un problème donné. Il dépendra de beaucoup de facteurs, y compris la complexité du problème, le nombre des données d'apprentissage, les valeurs des poids et le biais...etc.

Howard Demuth et Mark Beale présentent une comparaison entre les performances des différents algorithmes d'apprentissage par rétropropagation du gradient [HOWA00].

Le tableau suivant présente les algorithmes qui sont examinés et les abréviations utilisées pour les identifier :

Abréviation	Algorithme
LM	trainlm - Levenberg-Marquardt
BFG	trainbfg - BFGS Quasi-Newton
RP	trainrp - Resilient Backpropagation
SCG	trainscg - Scaled Conjugate Gradient
CGB	traincgb - Conjugate Gradient with Powell/Beale Restarts
CGF	traincgf - Fletcher-Powell Conjugate Gradient
CGP	traincgp - Polak-Ribiere Conjugate Gradient
OSS	trainoss - One-Step Secant
GDX	traingdx - Variable Learning Rate Backpropagation

Tableau 3.1- Différents Algorithmes d'apprentissage.

3.5.9 - Critères d'arrêt [PARIZ04] :

Plusieurs critères d'arrêts peuvent être utilisés avec l'algorithme de rétropropagation des erreurs. Le plus commun consiste à fixer un nombre maximum de périodes d'entraînement, ce qui fixe effectivement une limite supérieure sur la durée

de l'apprentissage. Ce critère est important car la rétropropagation des erreurs n'offre aucune garantie quant à la convergence de l'algorithme. Il peut arriver, par exemple, que le processus d'optimisation reste pris dans un minimum local. Sans un tel critère, l'algorithme pourrait ne jamais se terminer.

Un deuxième critère commun consiste à fixer une borne inférieure sur l'erreur quadratique moyenne, ou encore sur la racine carrée de cette erreur. Dépendant de l'application, il est parfois possible de fixer à priori un objectif à atteindre. Lorsque l'indice de performance choisi diminue en dessous de cet objectif, on considère simplement que le PMC a suffisamment bien appris ses données et on arrête l'apprentissage.

Les deux critères précédents sont utiles mais ils comportent aussi des limitations. Le critère relatif au nombre maximum de périodes d'entraînement n'est aucunement lié à la performance du réseau. Le critère relatif à l'erreur minimale obtenue mesure quant à lui un indice de performance; mais ce dernier peut engendrer un phénomène dit de sur-apprentissage qui n'est pas désirable dans la pratique, surtout si on ne possède pas une grande quantité de données d'apprentissage, ou si ces dernières ne sont pas de bonne qualité.

3.6 - Description des bancs d'essai :

Les signaux de simulation utilisés pour valider la capacité des réseaux de neurones artificiels à diagnostiquer les défauts dans les machines tournantes, sont enregistrés à partir de deux bancs d'essai suivants:

3.6.1-Banc d'essai d'engrenage (CETIM):

Le système à étudier, est un réducteur composé de deux roues (figure 3.4) [BOUI00], ce banc d'essai du CETIM 'Centre d'Etudes Techniques des Industries Mécaniques- France', a les caractéristiques suivantes :

- Le nombre de dents est respectivement 20 et 21 dents.
- La vitesse de rotation de l'arbre moteur est de **1000** tr / min .
- La fréquence d'engrènement $f_e = 330$ HZ .
- L'une des dents d'une roue de l'engrenage a été détériorée au cours de l'expérimentation.
-
-

- On a enregistré chaque jour et pendant 12 jours, un signal vibratoire sur le banc d'essai comportant 60160 échantillons avec une fréquence d'échantillonnage de 20 KHZ.

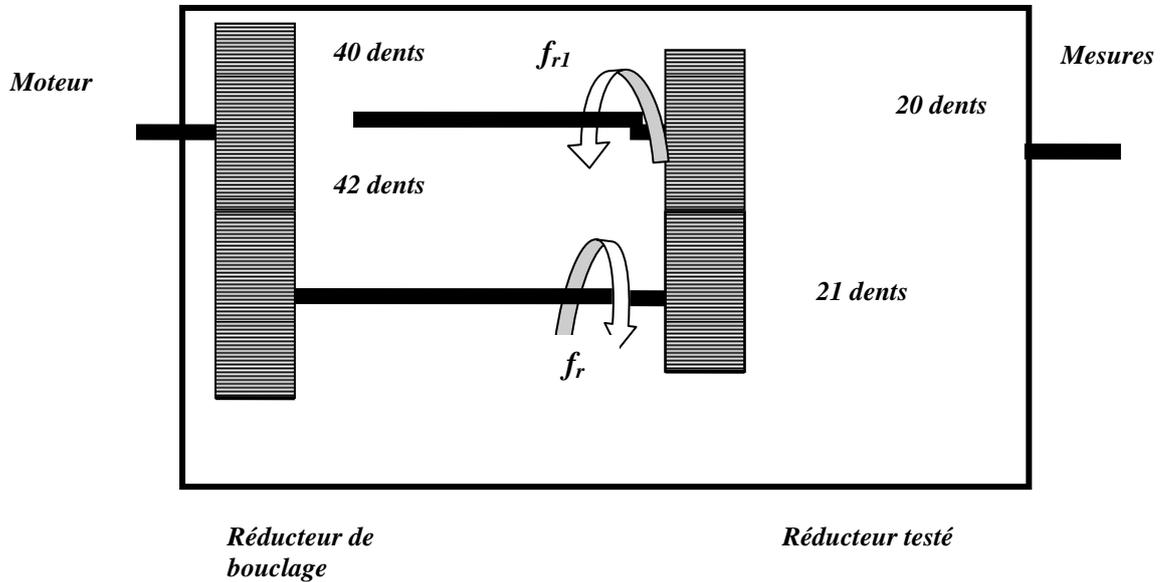


Figure3.4- Schéma du dispositif expérimental

3.6.1.1- Rapport d'expertise :

Le rapport d'expertise est présenté par le tableau 3.2 :

Jour	Observation
2	Premier jour d'acquisition, pas d'anomalie
3	Pas d'anomalie
4	// //
5	// //
6	// //
7	Écaillage sur le profil de la dent 1/2
8	Pas d'évolution
9	dent 1/2 pas d'évolution
10	Évolution de l'écaillage dent 15/16
11	// //
12	// //
13	Écaillage sur toute la largeur de la dent 15/16

Tableau 3.2 - Rapport d'expertise [ELBAD99b].

3.6.1.2- Représentation des signaux temporels:

La figure 3.5, nous montre les signaux temporels représentant 3 périodes de rotation. L'analyse des signaux temporels nous permet de visualiser l'apparition d'un défaut très marqué le douzième jour.

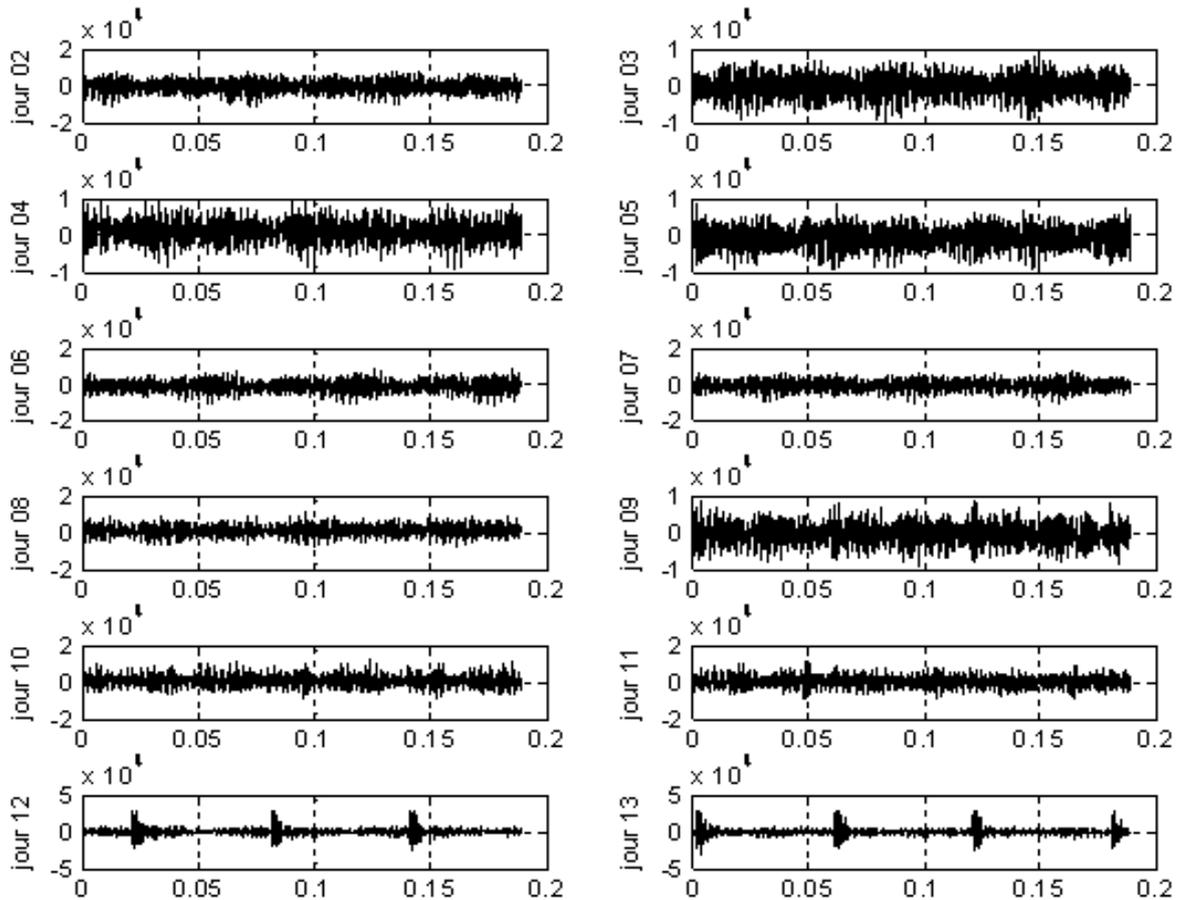


Fig3.10- Représentation des signaux temporels (CETIM).

3.6.2- Banc d'essai de roulements« the case western reserve university- bearing data centre»:

Le banc d'essai (voir figure 3.11) se compose d'un moteur "2 HP" (à gauche), d'un accouplement d'un transducer/encoder (centre), d'un dynamomètre (droit) et des circuits de commande (non montrés). Les roulements testés soutiennent l'arbre du moteur des deux cotés (coté entraînement et coté ventilateur). De simples défauts sous forme de points ont été créés sur les roulements d'essai, en utilisant l'usinage électro décharge [FEDA05].



Fig3.11- Banc d'essai de roulements.

L'accès à la description et aux caractéristiques des différents roulements, et à la base de données des différents essais est possible sur le site web du laboratoire "[www.eecs.cwru.edu/laboratory/bearing/welcome overview.htm](http://www.eecs.cwru.edu/laboratory/bearing/welcome%20overview.htm)".

Le tableau 3.3 représente les diamètres des défauts, leurs profondeurs et les marques de roulement utilisés.

Position du roulement	Emplacement du défaut	Diamètre [inches]	profondeur [inches]	Marque du roulement
coté entraînement	bague Intérieure	0.007	0.011	SKF
coté entraînement	bague Intérieure	0.014	0.011	SKF
coté entraînement	bague Intérieure	0.021	0.011	SKF
coté entraînement	bague Intérieure	0.028	0.050	NTN
coté entraînement	bague extérieure	0.007	0.011	SKF
coté entraînement	bague extérieure	0.014	0.011	SKF
coté entraînement	bague extérieure	0.021	0.011	SKF
coté entraînement	bague extérieure	0.040	0.050	NTN
coté entraînement	Bille	0.007	0.011	SKF
coté entraînement	Bille	0.014	0.011	SKF
coté entraînement	Bille	0.021	0.011	SKF
coté entraînement	Bille	0.028	0.150	NTN
coté ventilateur	bague Intérieure	0.007	0.011	SKF
coté ventilateur	bague Intérieure	0.014	0.011	SKF
coté ventilateur	bague Intérieure	0.021	0.011	SKF
coté ventilateur	bague extérieure	0.007	0.011	SKF
coté ventilateur	bague extérieure	0.014	0.011	SKF

coté ventilateur	bague extérieure	0.021	0.011	SKF
coté ventilateur	Bille	0.007	0.011	SKF
coté ventilateur	Bille	0.014	0.011	SKF
coté ventilateur	Bille	0.021	0.011	SKF

Tableau 3.3. Caractéristiques des défauts des roulements.

Les signaux ont été enregistrés à l'aide des accéléromètres, qui ont été fixés à la cage du moteur par aimant. Ces capteurs ont été placés à la position 12 heures "vertical", des deux cotés "accouplement et ventilateur" de la cage du moteur.

Les fréquences caractéristiques des éléments du roulement sont des multiples de la vitesse de rotation f_r :

- Fréquence de la bague intérieure: $f_{bi} = 5.415*(f_r)$
- Fréquence de la bague extérieure: $f_{be} = 5.584*(f_r)$
- Fréquence de la cage: $f_c = 0.3982*(f_r)$
- Fréquence de la bille: $f_b = 5.7135*(f_r)$

Tous les fichiers de données sont au format (*.mat). Chaque fichier contient trois données: signaux de vibrations des deux cotés et la vitesse de rotation du moteur. Pour notre étude nous avons considéré, seulement, es signaux du côté accouplement.

La figure 3.12 représente les signaux temporels recueillis du coté de l'accouplement, pour différents diamètres de défaut, réaliser sur la bague intérieure, pour un couple moteur de 3HP. (HP: Horse Power).

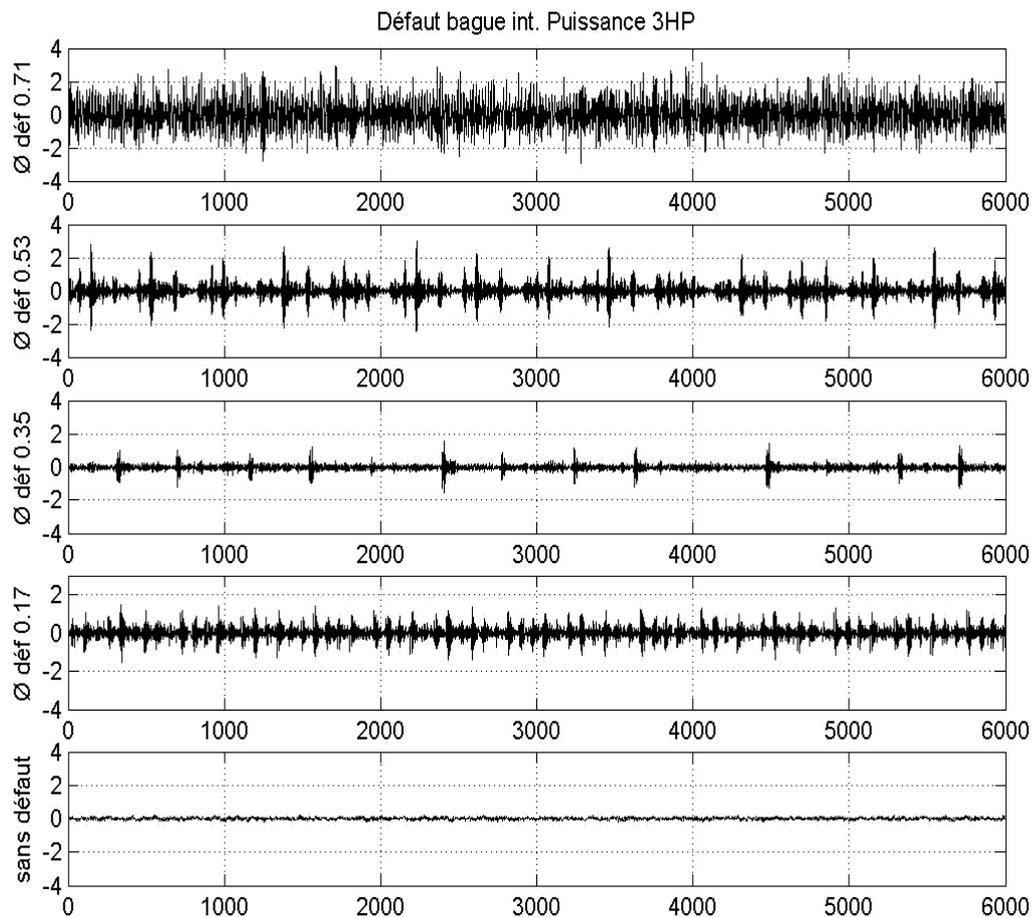


Fig3.12- Représentation des signaux temporels.

3.7 – Prétraitement et extraction des données :

Nous avons utilisé le logiciel *TRSIGN*^(*), pour traiter les signaux et extraire les caractéristiques correspondantes aux différents états des dispositifs expérimentaux. La figure 3.12 présente le menu principal du logiciel qui s'affiche à l'écran.

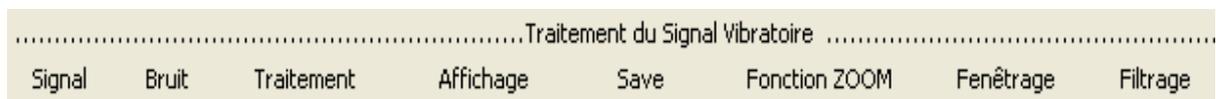


Fig3.13- Le menu principal du logiciel *TRSIGN*.

3.8– Classification par les RNAs:

Nous avons développé un programme sous environnement Matlab, qui offre la possibilité de classifier les différents états de santé des éléments des bancs d'essais étudiés, en utilisant les réseaux de neurones artificiels, et nous permet aussi de choisir et de varier les valeurs des paramètres du réseau utilisé.

^(*) : *TRSIGN*, est un logiciel du traitement de signal vibratoire développé au niveau de notre laboratoire de "LMPA".

Les programmes sont exécutés sur un micro-ordinateur Pentium4 avec une ram de 1G et une vitesse de 3.4GHZ.

3.9- Conclusion :

Ce chapitre a été consacré à la démarche expérimentale du diagnostic des machines tournantes, par la classification des différents signaux correspondant aux différents états de la machine en utilisant la méthode des réseaux de neurones artificiels. Nous avons expliqué cette démarche étape par étape. Nous avons trouvé que plusieurs paramètres participent dans le succès du diagnostic basé sur la méthode des réseaux de neurones artificiels. La difficulté de ce type de diagnostic, malgré les travaux et les études qui ont été faites dans ce sujet, reste toujours le choix des paramètres du réseau ; commençant par les éléments du vecteur forme, passant par l'architecture du réseau, en arrivant au vecteur de sortie. On note ici, qu'il n'y a pas de règles pour choisir les paramètres des réseaux de neurones artificiels, mais ces paramètres varient selon le problème étudié. Seule, l'expérience permet de répondre à cette question. Nous n'oublions pas la simplicité d'implémenter un réseau de neurones de type perceptron multicouche sous l'environnement du langage MATLAB. Il suffit de déterminer les différentes valeurs correspondant aux différents paramètres du réseau, puis d'exécuter les différentes commandes (Newff, train et sim).

Nous avons présenté dans ce chapitre, les deux bancs d'essai utilisés pour étudier les performances des réseaux de neurones artificiels utilisés comme un outil de diagnostic pour la classification automatique des défauts.

Dans ce sens, le chapitre suivant contiendra la discussion des différents résultats obtenus en étudiant l'influence des différents paramètres sur les performances du réseau de neurones artificiels.

CHAPITRE 4

RESULTATS ET DISCUSSION

CHAPITRE 4: RESULTATS ET DISCUSSION

4.1 - Introduction :

La variété et la diversité des paramètres influant les performances des RNAs (de type perceptron multicouches) rendent l'exploitation de ces derniers, en diagnostic des machines tournantes, très compliquée. Cependant, ces paramètres n'ont pas le même degré d'influence et d'importance sur les performances des RNAs.

Dans ce chapitre est discuté l'influence du choix de l'architecture et de la configuration du PMC (Perceptron multicouches) sur leurs performances de ce dernier. La première partie, traite l'effet du nombre maximal d'itérations, du nombre de neurones dans la couche cachée, de la valeur du taux d'apprentissage, de la valeur de l'erreur minimale, et du nombre des exemples d'apprentissage sur les performances du RNA. On a appliqué le PMC dans le diagnostic des engrenages (banc d'essai de CETIM). Le vecteur d'entrée contient les paramètres calculés à partir du modèle AR (Autorégressif) des différents signaux prélevés. Dans la deuxième partie, on étudie l'effet des mêmes paramètres cités précédemment, pour les roulements (banc d'essai de "the case western reserve university"). Le vecteur d'entrée contient des indicateurs issus du signal temporel et du spectre d'enveloppe.

4.2 – Diagnostic des défauts d'engrenage:

Les signaux que nous proposons d'étudier sont les signaux prélevés du banc d'essai de CETIM (Défauts d'engrenages). Le réseau de neurones utilisé dans cette application, est un réseau perceptron multicouche (PMC) avec un apprentissage de type rétropropagation du gradient. Mais, avant de commencer la classification, un prétraitement des signaux est nécessaire. Nous décrivons ci-dessous les différentes étapes traitées.

4.2.1- Constitution de la base des données:

Nous avons signalé précédemment, que les RNA nécessitent une grande base de données pour faire l'apprentissage et le test. Nous avons seulement pour ce banc, 12

signaux correspondant aux jours d'enregistrement, ce qui ne nous permet pas de construire la base de données demandée.

4.2.1.1- Préparation des données:

Pour avoir suffisamment de données, nous avons divisé chaque signal en K_t tranches de N_M points, le nombre d'échantillons pour chaque tranche doit être suffisamment grand, pour couvrir un nombre suffisant du plus grand cycle correspondant à la plus grande période (la plus petite vitesse de rotation).

Nous disposons des signaux de longueur de 60160 points. Le moteur tourne à une vitesse de 1000 tr/min, le rapport de réduction était 20/21.

$$N_r = 1000 \text{ tr/min} \Rightarrow fr_1 = \frac{1000}{60} = 16,67 \text{ Hz}$$

$$fr_2 = \frac{20}{21} \cdot fr_1 = 15,88 \text{ Hz}.$$

Les périodes T_1 et T_2 , respectivement de la roue 1 et 2 sont :

$$T_1 = \frac{1}{fr_1} = \frac{6}{100} = 0,06 \text{ s}$$

$$\text{et } T_2 = \frac{1}{fr_2} = \frac{6.3}{100} = 0.063 \text{ s}$$

La fréquence d'échantillonnage est de : $fe = 20 \text{ kHz} = 20000 \text{ Hz}$

La période d'échantillonnage sera : $\Delta t = \frac{1}{fe} = 5 \cdot 10^{-5} \text{ s}$.

Le nombre de points Np_1 et Np_2 par période respectivement de la roue 1 et 2 est :

$$Np_1 = \frac{T_1}{\Delta t} = \frac{fe}{fr_1} = \frac{20000}{16.67} = 1200 \text{ pts}$$

$$\text{et } Np_2 = \frac{T_2}{\Delta t} = \frac{fe}{fr_2} = 1260 \text{ pts}$$

La plus grande période est $T_2 = 0.063 \text{ s}$

Si nous prenons un coefficient de 5. La longueur nécessaire de la tranche sera :

$$Tt = 0.063 * 5 = 0.315 \text{ s}.$$

Le nombre d'échantillons nécessaire par tranche est de :

$$Ne = \frac{Tt}{\Delta t} = 1260 * 5 = 6300 \text{ pts}$$

Si nous prenons pour N_M une valeur de 2^n supérieure à N_e ,
 $N_M = 2^{13} = 8192$ échantillons.

Alors le nombre de tranche dans le cas d'un recouvrement 50 %, soit $s_1 = \frac{N_M}{2}$, est de :

$$K_t = \frac{N_s - s_1}{N_M - s_1} \cong 13 \text{ tranches} .$$

Donc, Nous avons obtenu 156 signaux, pour toute la durée de l'expérience (12 jours).
 C'est-à-dire 12 jours x 13 tranches.

4.2.1.2- Prétraitement de données :

a- Le filtrage:

Dans Boulanger et al. [BOUL03], on indique qu'il existe toujours un effet de masque qui entache fortement la sensibilité des indicateurs "larges bandes" lorsque l'énergie vibratoire, introduite par l'apparition d'un défaut, n'est pas prépondérante devant celle introduite par l'ensemble des forces dynamiques générées, par le fonctionnement de la machine et des défauts déjà existants. Pour notre travail, les indicateurs sont donc calculés dans certains nombres de bandes fréquentielles ni trop larges ni trop étroites et définies en fonction des caractéristiques cinématiques de la machine.

Nous avons choisi, après une analyse préliminaire [FELK007], de calculer les indicateurs temporels dans de différentes bandes latérales de largeur 167 Hz autour des fréquences d'engrènement; soit 10 raies latérales (dix fréquences de rotation) autour des 3 principales harmoniques ($1.f_e=333.33 \text{ Hz}$, $2.f_e=666.66 \text{ Hz}$, et $3.f_e=999.99 \text{ Hz}$). Les bandes de calcul sont alors les suivantes : [250-417 Hz], [583-750 Hz] et [916-1083 Hz].

Le signal de chaque tranche a donc été centré puis filtré par un filtre passe-bande. Le filtre utilisé est à réponse impulsionnelle finie à phase minimale.

b- La décimation:

Pour Chaque jour, nous disposons de 13 tranches de signaux de $N_M=8192$ points, échantillonnées à 20000 Hz . Par le filtrage ci-dessus, nous avons limité notre étude à une fréquence maximum de 1083 Hz . Les tranches des signaux sont, donc, suréchantillonnées. Nous proposons de décimer ces signaux par un facteur 4. Nous obtenons ainsi des signaux de 2048 points échantillonnés à une fréquence $f_e=20000/4=5000 \text{ Hz}$. Nous pouvons remarquer que la demi fréquence

d'échantillonnage de Nyquist reste supérieure à la fréquence maximale du signal analysé. Le théorème de Shannon ($f_e \geq 2 \cdot f_{\max}$) est donc bien vérifié. Le vecteur décimé est 4 fois plus court.

4.2.1.3- Calcul des paramètres AR :

Basés sur des travaux antérieurs de Felkaoui et al, [FELK95], qui ont développé un algorithme d'estimation des paramètres. Les simulations numériques ont montré que l'algorithme proposé donne une bonne poursuite des variations de paramètres même en présence de bruit. On a choisi de calculer 60 paramètres à partir du modèle AR.

4.2.1.4- Constitution du vecteur forme (Entrée du RNA):

Les données que l'on doit classer et traiter sont stockées dans un tableau de type observations/variables de la forme suivante :

$$BD = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1M} \\ x_{21} & x_{22} & \dots & x_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \dots & x_{NM} \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix}$$

où $\begin{cases} N \text{ est le nombre d'observations} \\ M \text{ est le nombre d'indicateurs} \end{cases}$

On a alors :

M variables, représentées en colonnes, et N observations représentées en lignes qui seront divisées en N_c classes (w_1, w_2, \dots, w_{N_c}) ou modes de fonctionnement. Chaque classe w_i possède N_i observations. Pour notre travail:

- $N = 156$ (13 tranches par 12 signaux)
- $M = 60$ (indicateurs) \longrightarrow L'entrée du RNA.
- $N_c = 2$ classes \longrightarrow La sortie du RNA.

4.2.1.5- Choix des classes (Sortie du RNA):

Le vecteur de sortie contient les différentes classes correspondant aux différents états de fonctionnement du dispositif expérimental. En absence du défaut l'étiquette est 1, dans le cas contraire l'étiquette est 0.

Le rapport de l'expertise montre l'apparition d'un début d'écaillage sur la dent 15/16 dès le 7^{ème} jour. On considère la variante suivante:

- Du 2^{ème} jour au 11^{ème} jour : étiquette 1 (fonctionnement sain).
- Du 12^{ème} jour au 13^{ème} jour: étiquette 0(fonctionnement défaillant).

4.2.1.6- Normalisation des données:

Afin d'éviter qu'un paramètre ne prenne trop d'importance par rapport aux autres, tout simplement parce qu'il varie dans de plus grandes proportions (plage), on procède généralement à une normalisation des données en divisant les composantes de chaque paramètre par son écart-type et en les centrant (Variable réduite et centrée).

$$x_{ij} = \frac{x_{ij} - m_j}{\sigma_j} \quad (4.1)$$

Où σ_j représente l'écart- type du j^{ème} paramètre et m_j sa moyenne.

La base de données obtenue sera divisée en deux parties: Une Base d'Apprentissage (BA) qui entraîne le réseau, et une Base de Test (BT) des échantillons restant sur laquelle on mesure les performances du réseau.

La matrice d'apprentissage BA contient :

- de la ligne 1 à 70 : la classe 1 (7 tranches x 10 jours).
- de la ligne 71 à 84 : la classe 2 (7 tranches x 2 jours).

La matrice de test BT contient alors :

- de la ligne 1 à 60 : la classe 1 (6 tranches x 10 jours).
- de la ligne 61 à 72 : la classe 2. (6 tranches x 2 jours).

4.2.2. Exploitation du système du diagnostic basé sur les RNA :

Afin de pouvoir exploiter les capacités de classification des RNAs, et les adapter à notre application, nous avons développé un programme sous environnement MATLAB. Il nous permet de faire la classification et le test des différents états du dispositif expérimental, avec les différentes structures et les différents paramètres du réseau.

4.2.2.1- Configuration du RNA:

En se basant sur une analyse préliminaire et sur la synthèse bibliographique du chapitre2, nous avons utilisé un réseau perceptron multicouches avec la configuration suivante:

- Nombre d'individus total 156 (84 pour l'apprentissage, 72 pour le test) ;
-

- 60 neurones dans la couche d'entrée, correspondant aux éléments du vecteur forme.
- Une seule couche cachée.
- Un neurone dans la couche de sortie correspondant au vecteur de sortie qui prend l'une des deux valeurs :
 - **1** dans le cas de l'absence du défaut (état sain).
 - **0** dans le cas de la présence du défaut (état défectueux).
- La fonction de transfert tansig (Tangente sigmoïde) pour la couche d'entrée, et la fonction purelin (Linéaire) pour la couche cachée.
- L'apprentissage se fait par l'algorithme de la rétropropagation du gradient basé sur l'algorithme de Levenberg Marquardt (trainlm).
- Fonction de performance : mse (Mean square error).

Les valeurs du nombre de neurones dans la couche cachée et du nombre maximal d'itérations, de l'erreur minimale, et du taux d'apprentissage, ainsi que du nombre des exemples d'apprentissage ne sont pas définis, car l'objectif de cette partie du travail est d'étudier leurs effets sur les performances du RNA .

Le critère qui nous permet d'évaluer les performances du réseau est le pourcentage de classification (taux de succès) T.

$$T = \frac{N_{Bc}}{N} \times 100$$

Avec

N_{Bc} : Nombre d'observations bien classées.

N : Nombre total d'observations.

Les programmes sont exécutés sur un micro-ordinateur Pentium4 avec une RAM de 1G et une vitesse de 3.4G.

4.2.2.2- Effet du nombre d'itérations maximal:

Le nombre d'itérations maximal est l'un des critères d'arrêt du processus d'apprentissage, c'est-à-dire quand le réseau atteint le nombre fixé par l'utilisateur, l'apprentissage s'arrête. Dans cette section, nous étudions les performances du réseau en fonction du nombre d'itérations. Nous varions le nombre du 1 à 1000 itérations, et nous faisons plusieurs essais (Total de 10 essais) pour chaque nombre maximal d'itérations. Puis, nous calculons la moyenne et l'écart type du taux de classification

correspondant à chaque nombre maximal d'itérations. Les résultats obtenus sont présentés dans le **tableau 4.1**:

Nombre d'itérations	Pourcentage des bien classés (Apprentissage)	Ecart- type Apprentissage	Pourcentage des bien classés (Test)	Ecart- type Test
1	81.9	0.93	81.66	1.09
5	83.69	1.49	82.77	0.78
10	86.78	3.94	85.13	3.21
50	96.66	2.08	92.35	2.71
100	99.16	1.13	96.24	1.85
150	100	0	95.83	4.44
200	100	0	97.77	1.63
300	100	0	96.80	2.27
400	100	0	95.69	3.03
500	100	0	95.96	2.40
1000	100	0	93.88	5.90

Tableau 4.1 – Effet du nombre maximal d'itérations.

La figure 4.1 présente les courbes de la variation des taux d'apprentissage et du test en fonction du nombre d'itérations.

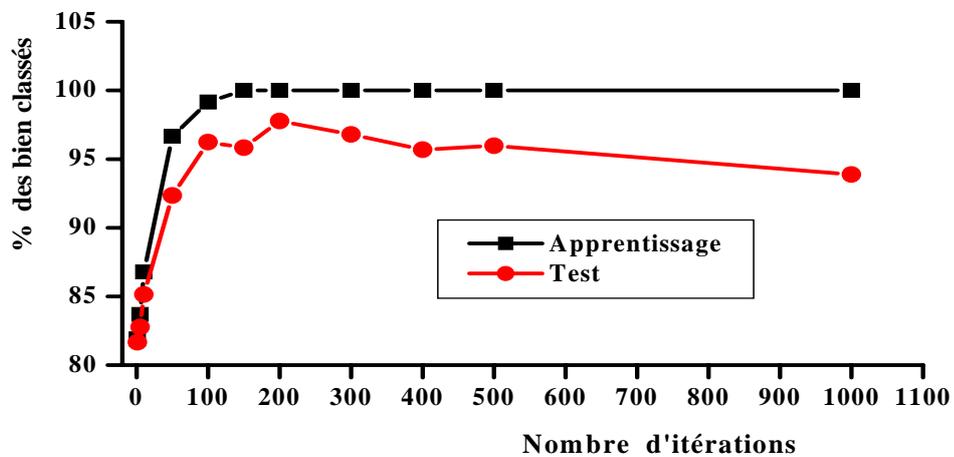


Fig4.1 - Effet du nombre maximal d'itérations.

- Discussion:

La figure 4. 1 présente l'évolution du pourcentage des observations bien classées en fonction du nombre d'itérations. On constate que la performance en apprentissage s'améliorent au fur et à mesure que le nombre d'itérations augmente et atteint 100% à l'itération 150. Mais généralement, l'objectif est d'être performant en généralisation c'est-à-dire avec des données non- apprises pour éviter un sous-apprentissage ou un sur- apprentissage. Le nombre d'itérations est donc fixé par la courbe de test, où nous avons remarqué que les performances commencent à se dégrader après un nombre d'itérations égale à 200. Le RNAs donne le meilleur taux de reconnaissance (classification) optimal des données test de 97.7 % avec un écart-type faible de 1.63, calculé sur 10 essais ce qui implique une petite dispersion des valeurs autour de la moyenne et donc, une confiance élevée sur le taux d'observations bien prédites (une petite marge d'erreur), voir figure 4.2.

Nous avons donc choisi, un nombre optimal d'itérations de 200 pour la suite des essais.

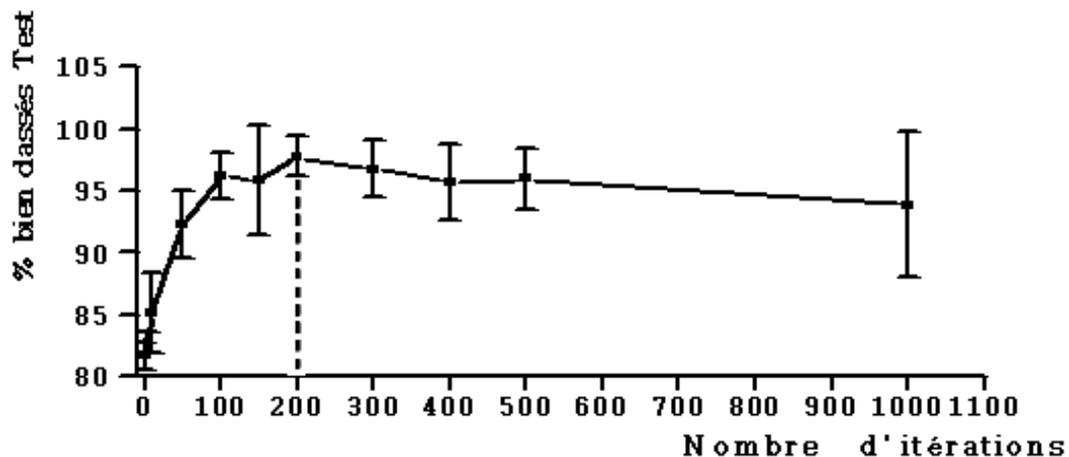


Fig4.2- Pourcentage des observations bien classés test.

4.2.2.3- Effet du nombre des neurones dans la couche cachée:

Le nombre de neurones dans la couche cachée est le paramètre le plus important en ce qui concerne les possibilités d'approximation d'un réseau de neurones artificiels, car il influe directement sur la convergence de l'algorithme de la rétropropagation du gradient. Dans cette section, nous allons étudier l'influence du choix de ce paramètre sur les performances de classification d'un perceptron multicouche. Pour cela, nous utilisons la même configuration présentée dans la section 4.2.2.1, en plus d'un nombre maximal d'itérations de 200. Nous varions le nombre des neurones dans la couche cachée de 1 à 20. Nous faisons plusieurs essais (Total de 10

essais) pour chaque nombre de neurones cachés. Les résultats obtenus sont présentés dans le tableau 4.2.

Nombre des neurones cachés	Pourcentage des bien classés (Apprentissage)	Ecart- type Apprentissage	Pourcentage des bien classés (Test)	Ecart- type Test
1	94,75	4,38	92,79	5,43
2	98,09	,211	95,13	3,16
3	98,68	1,42	96,29	2,03
4	99,64	0,87	96,94	2,25
5	100	0	98,89	1,2
6	100	0	97,50	2,91
7	100	0	97,37	2,57
8	100	0	97,3	2,03
10	100	0	97,22	2,10
15	100	0	95,55	2,68
20	100	0	94,02	3,98

Tableau 4.2- Effet du nombre de neurones dans la couche cachée sur le pourcentage d'observations bien cassées.

La figure 4.3 présente les courbes de la variation des observations bien classées (d'apprentissage et du test) en fonction du nombre de neurones dans la couche cachée.

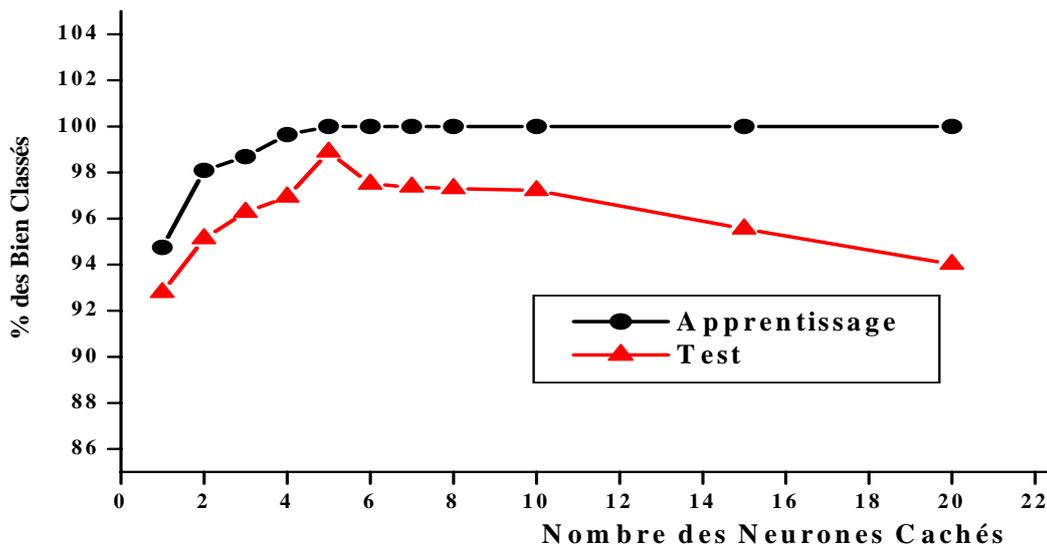


Fig.

4.3- Effet du nombre des neurones dans la couche cachée sur le pourcentage d'observations bien cassées.

- Discussion:

La figure 4.3 présente l'évolution du taux des observations bien classés (apprentissage et test) en fonction du nombre de neurones dans la couche cachée. On remarque que la performance de l'apprentissage augmente vis-à-vis du nombre de neurones dans la couche cachée, jusqu'à la valeur maximale (100%), qui correspond à un nombre de neurones égal à 5. Pour juger les performances du RNAs en généralisation, c'est-à-dire avec des données non- apprises et surtout pour éviter la saturation du réseau (la convergence du réseau devient très lente), on doit se référer à la courbe des biens classés du test. Le nombre de neurones dans la couche cachée est ainsi fixé par cette courbe où nous avons remarqué que la performance augmente jusqu'à atteindre la valeur maximale (98,89 %), qui correspond à un nombre de 5 neurones. Au-delà de ce maximum, elle commence à se dégrader.

Nous remarquons que l'écart type (dispersion) le plus faible, calculé sur 10 essais, correspond au nombre de neurones égal 5, voir figure 4.4. Nous avons donc choisi, un nombre optimal de neurones dans la couche cachée de 5 pour la suite des essais.

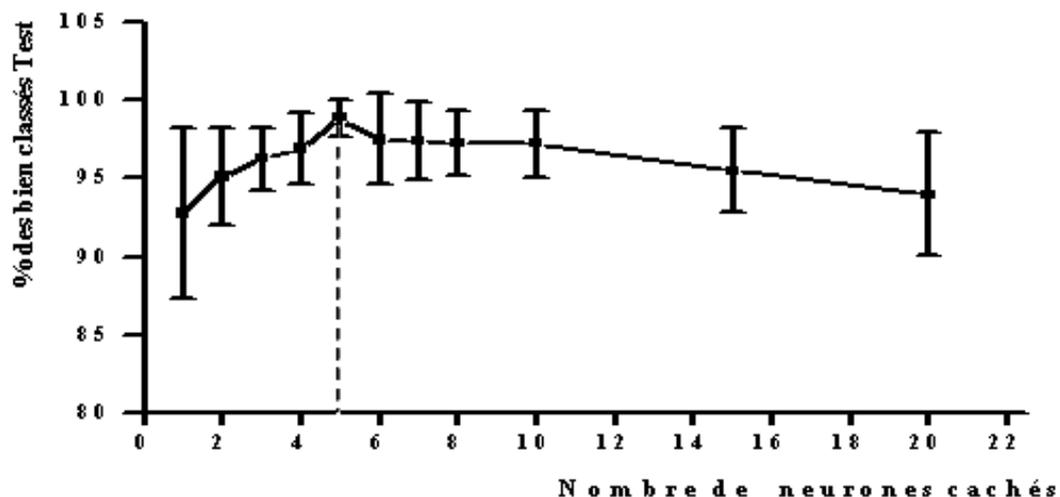


Fig4.4- Pourcentage des bien classés du test.

4.2.2.4- Effet de la valeur du taux d'apprentissage :

La valeur du taux d'apprentissage (η) est incluse dans l'étape de rétropropagation de l'erreur, donc elle influe directement sur l'ajustement (correction) des valeurs des poids synaptiques, cette valeur varie généralement entre 0 et 1.

Dans cette section, nous proposons d'étudier l'effet des différentes valeurs du taux d'apprentissage sur la performance de la classification du perceptron

multicouches. En faisant varier ces valeurs dans un intervalle de $[0.1, 0.9]$, nous avons gardé la même configuration présentée dans les sections précédentes (Nombre maximal d'itérations égal 200 et nombre de neurones dans la couche cachée égal 5). Nous faisons 10 essais pour chaque valeur de η et nous calculons la moyenne de pourcentage des bien classés pour l'apprentissage et pour le test. Les résultats obtenus sont présentés sur le **tableau 4.3**.

Taux d'apprentissage	Pourcentage des bien classés (Apprentissage)	Ecart- type Apprentissage	Pourcentage des bien classés (Test)	Ecart- type Test
0.1	99.76	0.50	95.13	1.63
0.2	99.88	0.38	95.82	3.07
0.3	99.88	0.38	95.83	2.78
0.4	100	0	95.27	2.19
0.5	99.88	0.38	96.30	1.66
0.6	100	0	96.10	2.25
0.7	100	0	96.74	2.58
0.8	100	0	98.61	1.6
0.9	99.86	0.43	96.52	3.15

Tableau 4.3- Effet du taux d'apprentissage.

- Discussion:

L'effet de la valeur du taux d'apprentissage (η) est schématisé sur la figure 4.5, qui représente l'allure des pourcentages des bien classés d'apprentissage et du test pour différentes valeurs de η .

L'analyse de ces courbes montre que le pourcentage des biens classés de l'apprentissage présente généralement une droite avec l'augmentation de η . Tandis que le pourcentage des biens classés de test présente une croissance avec l'augmentation de η . Les performances du réseau deviennent maximales pour $\eta = 0.8$. Ensuite, on remarque une diminution de cette valeur.

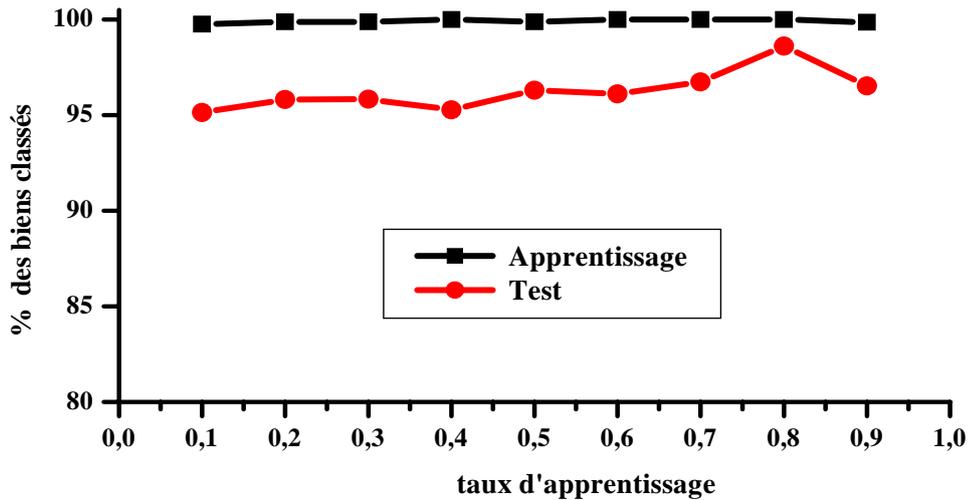


Fig.4.5- Effet du taux d'apprentissage sur le taux d'observations bien classées.

L'écart-type le plus faible calculé sur 10 essais, correspond également à une valeur de taux d'apprentissage égal 0.8 (voir figure 4.6). Alors, on peut dire que la valeur optimale du taux d'apprentissage, dans notre application est **0.8**.

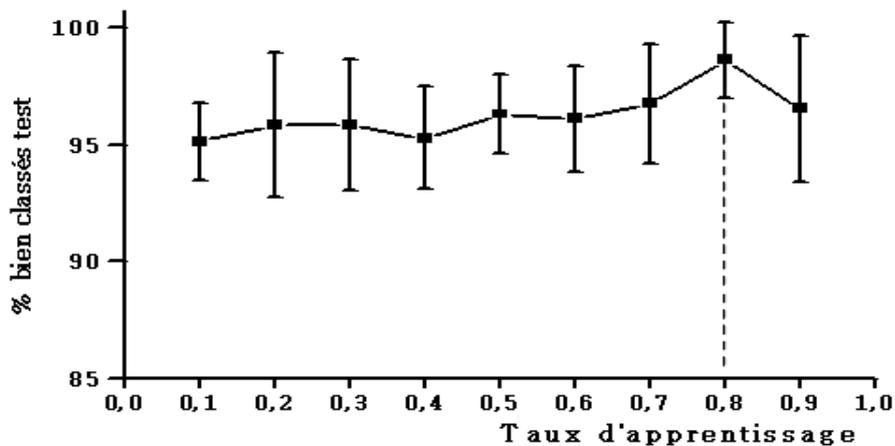


Fig4.6- Pourcentage d'observations bien classées du test

4.2.2.5- Effet de la valeur de l'erreur minimale (fonction de performance) :

La valeur de l'erreur minimale (mse^*) est l'un des **critères d'arrêt** du processus d'apprentissage, elle est sa borne inférieure. Elle influe directement sur la convergence du réseau. Dans cette partie du travail, nous cherchons la valeur optimale de l'erreur et son effet sur la performance du PMC. Nous varions sa valeur dans un intervalle de

(*) : mse , mean square error (erreur quadratique moyenne).

$[10^{-1}, 10^{-8}]$, (avec un pas de 10^{-1}). Nous faisons plusieurs essais (Total de 10 essais) pour chaque valeur de l'erreur. Puis, nous calculons la moyenne et l'écart type du taux de la classification.

Le Tableau 4.4 représente les différents résultats obtenus:

Erreur MSE	Pourcentage des bien classés (Apprentissage)	Ecart- type Apprentissage	Pourcentage des bien classés (Test)	Ecart- type Test
10^{-8}	100	0	95,96	2,01
10^{-7}	100	0	96,52	2,0
10^{-6}	100	0	96,10	2,15
10^{-5}	100	0	96,24	2,17
10^{-4}	100	0	96,38	2,38
10^{-3}	100	0	96,94	1,94
10^{-2}	100	0	96,66	2,19
10^{-1}	86,74	2,967	84,44	2,25

Tableau 4.4- Effet de la valeur minimale de l'erreur.

- Discussion:

La figure 4.7 présente l'évolution du taux des observations bien classés (apprentissage et test) en fonction de la valeur de l'erreur minimale. On constate que la performance de l'apprentissage présente généralement une droite avec l'augmentation de mse . Pour juger les performances du RNAs en généralisation, c'est-à-dire avec des données non- apprises, on doit se référer à la courbe des biens classés du test. La valeur minimale de mse , pour laquelle la performance atteint la valeur maximale (96,94 %) correspond à une valeur de 10^{-3} .

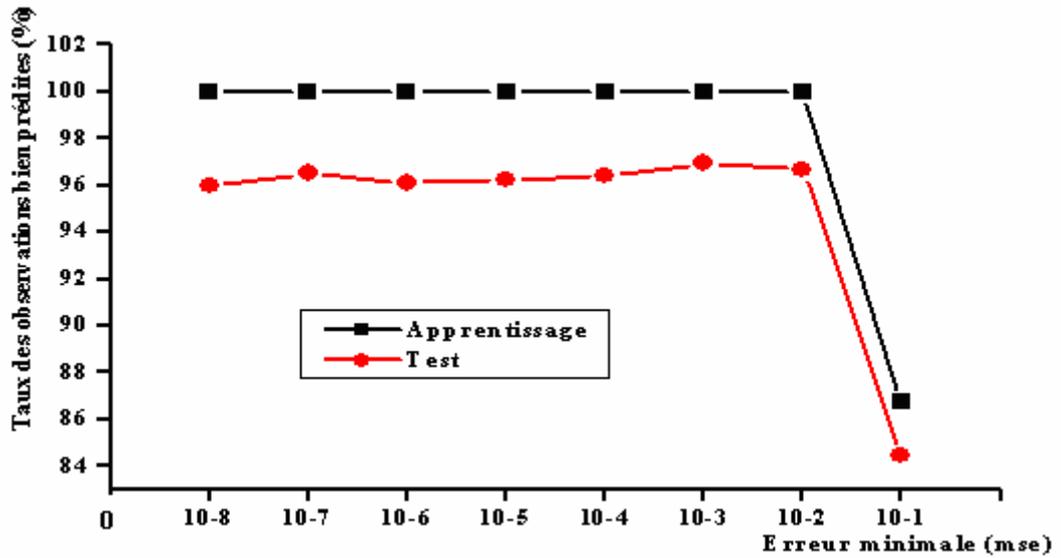


Fig4.7- Effet de l'erreur minimale sur le taux des observations bien classées.

Nous remarquons que l'écart type (dispersion) le plus faible, calculé sur 10 essais, correspond à une valeur de 10^{-3} , voir figure 4.8. Nous avons donc choisi, la valeur optimale de l'erreur minimale de 10^{-3} .

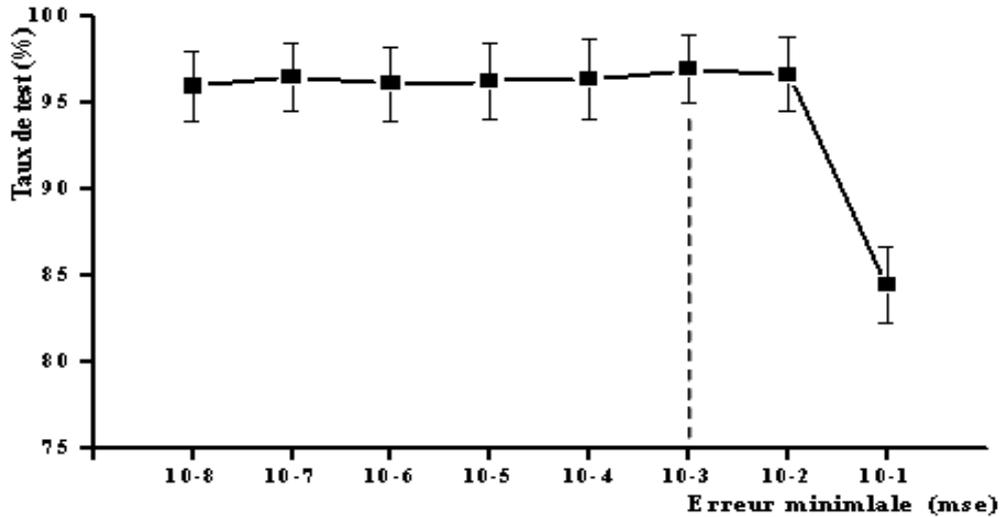


Fig4.8- Taux d'observations bien classées du test

4.2.2.6- Effet du nombre des exemples d'apprentissage :

L'apprentissage des RNAs est l'étape la plus importante dans la configuration des réseaux de neurones. Puisque, la création d'une règle de décision (absence ou présence de défauts) est basée sur la capacité de généralisation du réseau.

Dans cette section, nous allons étudier l'effet du nombre des exemples d'apprentissage sur les performances du RNAs. Nous commençons par un exemple

pour chaque jour, jusqu'au le nombre maximal offert par le prétraitement des signaux vibratoires (7 exemples pour chaque jour). C'est-à-dire, le nombre des lignes de la matrice d'apprentissage varie entre 12 et 84. Nous faisons 10 essais pour chaque exemple d'apprentissage. Ensuite, nous calculons la moyenne de pourcentage des bien classés pour l'apprentissage et pour le test. Le nombre des exemples de test (observations par jour) est fixé à 5. Pour la configuration du réseau, nous utilisons les valeurs optimales trouvées précédemment.

Les résultats obtenus sont présentés sur le *tableau 4.5*.

Nombre des exemples d'apprentissage pour chaque jour	Pourcentage des bien classés (Apprentissage)	Ecart- type Apprentissage	Pourcentage des bien classés (Test)	Ecart- type Test
1	100	0	80	0
2	100	0	80	0
3	100	0	78,87	1,74
4	100	0	80	0,86
5	100	0	79,49	0,80
6	100	0	90,99	4,52
7	100	0	92,46	1,64

Tableau 4.5- *Effet du nombre des exemples d'apprentissage.*

Discussion:

Les courbes représentées sur la figure **4.7** montrent l'évolution du taux de classification des observations bien prédites test en fonction du nombre des exemples d'apprentissage. En effet, on remarque que la performance du réseau est moyenne (80 %) pour un faible nombre d'exemples d'apprentissage; mais après l'augmentation des exemples d'apprentissage (>5), on aura une performance élevée (90%). Alors, pour qu'un réseau puisse reconnaître et classer tous les cas possible correspondant au fonctionnement de la machine, c'est-à-dire, créer une règle de généralisation (poids synaptiques optimaux), il faut que la base d'apprentissage recouvre tous les cas possibles du fonctionnement (état de santé) de la machine.

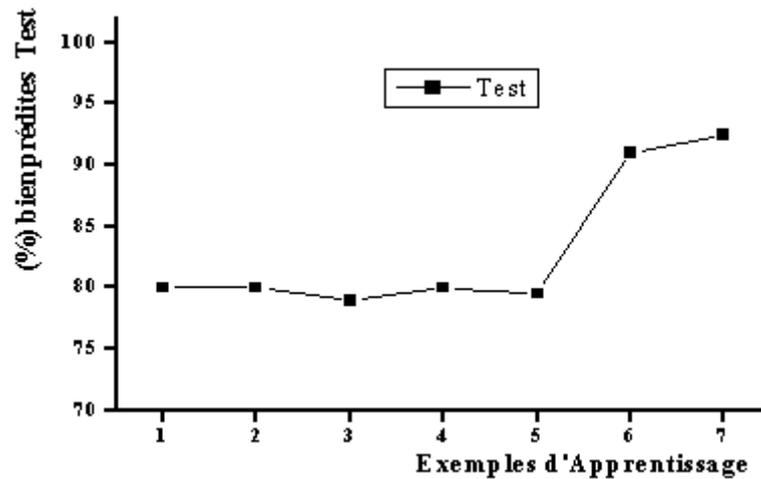


Fig4.7- Pourcentage des bien classés de test

4.3 – Diagnostic des défauts de roulements:

Dans cette partie du travail, les signaux que nous avons utilisé sont prélevés du banc d'essai Data centre bearing (défauts de roulements). Le réseau de neurones utilisé dans cette application, est un réseau perceptron multicouche avec un apprentissage de type rétropropagation du gradient.

Les signaux que nous proposons d'étudier sont enregistrés pour les défauts de la bague intérieure dans les conditions suivantes:

- L'arbre tourne à une vitesse de 1730 tr/min.
- La puissance (couples résistants) est de 3HP.
- Quatre diamètres de défauts: 0.17, 0.35, 0.53 et 0.71mm, ont été créés sur la bague intérieure.
- Les signaux sont prélevés du côté accouplement.

Avant de commencer la classification, on procède à un prétraitement des signaux.

4.3.1- Constitution de la base des données:

La base de données disponible contient cinq signaux; un signal sans défauts et les quatre autres avec défauts. Comme, les RNA nécessitent une grande base de données pour faire l'apprentissage et le test, et pour avoir suffisamment de données, nous avons divisé chaque signal en K_t tranches de N_M points à condition que le nombre d'échantillons pour chaque tranche doit être suffisamment grand, pour couvrir un nombre suffisant du plus grand cycle correspondant à la plus grande période (la plus petite vitesse de rotation).

Nous disposons des signaux de longueur de 120801 points. L'arbre tourne avec une vitesse de 1730 rpm.

$$f_{r_{\text{arbre}}} = \frac{1730}{60} = 28,83\text{Hz}$$

On a quatre fréquences caractéristiques correspondants aux différents éléments du roulement (cage, bagues extérieure et intérieure et les éléments roulants). La plus grande période correspond à la plus petite fréquence.

Dans notre cas, on doit prendre la fréquence de la cage:

$$f_{\text{cage}} = 0.9382 \times 28.83 = 11.48\text{HZ}$$

- La période :

$$T_{\text{cage}} = \frac{1}{f_{\text{cage}}}$$

- La fréquence d'échantillonnage :

$$f_e = 12\text{HZ} = 12000\text{HZ}$$

La période d'échantillonnage:

$$\Delta t = \frac{1}{f_e}$$

Le nombre de points par période:

$$N_p = \frac{T_{\text{cage}}}{\Delta T} = \frac{f_e}{f_{\text{cage}}} = \frac{12000}{11.48} = 1045.29 \text{pts} = 1046 \text{pts}.$$

Si nous prenons un coefficient de 5. La longueur nécessaire pour chaque tranche sera:

$$N_{\text{cage}} = N_p \times 5 = 5230 \text{pts}.$$

Si on prend pour N_m une valeur de 2^n supérieure à N_{cage} ,

$$N_m = 2^{13} = 8192 \text{ échantillons}$$

Alors le nombre de tranches dans le cas d'un recouvrement de 50%, soit

$$S1 = \frac{N_m}{2}$$

Alors:

$$K_t = \frac{N_s - S_1}{N_m - S_1} = \frac{120801 - 4096}{8192 - 4096} \cong 28 \text{ Tranches}.$$

Donc nous avons obtenu 140 signaux ; 28 signaux pour chaque classe.

4.3.1.1- Prétraitement et calcul des indicateurs:

Un prétraitement des signaux est nécessaire pour éliminer toute sorte d'information inutile, et de faciliter la tâche d'extraction des indicateurs de surveillance les plus pertinents formants la base de données.

Nous avons choisi de calculer les indicateurs suivants: la valeur efficace, le facteur de crête, la valeur crête à crête et le kurtosis, ainsi que les énergies issues du spectre enveloppe. Ces indicateurs peuvent être des bons indicateurs de surveillance dans le cas des roulements [BREN02], [PING00], [SUN98].

Après une analyse préliminaire [FEDA05], nous avons choisi de calculer ces indicateurs comme suit:

- Indicateurs temporels:

Les indicateurs temporels sont calculés en 5 bandes fréquentielles issues de la bande d'analyse " 6 KHZ", en 4 sous bandes adjacentes de largeur 1500 HZ, plus la bande totale. Les bandes de calcul sont alors: [1-1500HZ], [1500-3000HZ], [3000-4500HZ], [4500-6000HZ] et [1-6000HZ]. le signal de chaque tranche a donc été centré puis filtré par un filtre passe-bande.

- Indicateurs fréquentiels:

Les indicateurs fréquentiels sont calculés en 5 bandes fréquentielles de largeur de 1000HZ, issues du spectre d'enveloppe plus la bande totale. Alors les bandes de calcul sont [1-1000HZ], [1000-2000HZ], [2000-3000HZ], [3000-4000HZ], [4000-5000HZ] et [1000-6000HZ].

4.3.1.2- Constitution du vecteur forme (Entrée du RNA):

Le vecteur forme est constitué des indicateurs temporels et fréquentiels décrits précédemment. Les données que l'on doit classer et traiter sont stockées dans un tableau de type observations/variables de la forme suivante :

$$BD = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1M} \\ x_{21} & x_{22} & \cdots & x_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{NM} \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix}$$

où $\begin{cases} N \text{ est le nombre d'observations} \\ M \text{ est le nombre d'indicateurs} \end{cases}$

On a alors :

M variables, représentées en colonnes et N observations, représentées en lignes qui seront divisés en N_c classes (w_1, w_2, \dots, w_{N_c}) ou modes de fonctionnement. Chaque classe w_i possède N_i observations. Pour notre travail (cas de roulements):

- $N = 140$ (28 tranches par 5 signaux)
- $M = 26$ (indicateurs) \longrightarrow L'entrée du RNA.
- $N_c = 5$ classes \longrightarrow La sortie du RNA.

4.3.1.3- Choix des classes (Sortie du RNA):

Le vecteur de sortie du réseau contient les différentes classes correspondant aux différents états de fonctionnement du dispositif expérimental. Nous avons choisi cinq classes, chacune d'elles correspond à un diamètre du défaut. Le Tableau 4.6 représente l'étiquetage des différentes classes étudiées:

<i>Classes</i>	<i>Diamètre du défaut</i>	<i>Etiquette</i>
1	<i>Sans défaut</i>	<i>10000</i>
2	<i>0.17mm</i>	<i>01000</i>
3	<i>0.35mm</i>	<i>00100</i>
4	<i>0.53mm</i>	<i>00010</i>
5	<i>0.71mm</i>	<i>00001</i>

Tableau 4.6-*Etiquetage des classes.*

4.3.1.4- Normalisation des données:

Afin d'améliorer les performances du PMC, il est préférable de normaliser les données du vecteur forme, en utilisant l'équation (4.1). La normalisation sous MATLAB se fait par la commande `Prestd`.

On divise la base de données obtenue en deux parties: Une base d'apprentissage (BA) qui entraîne le réseau, et une base de test (BT) des échantillons restant sur laquelle on mesure les performances du réseau.

4.3.2 - Configuration du RNA:

On a utilisé un réseau perceptron multicouches avec la configuration suivante:

- Nombre d'individus total 140 (90 pour l'apprentissage, 50 pour le test).

- 26 neurones dans la couche d'entrée, correspondant aux éléments du vecteur forme.
- Une seule couche cachée.
- 5 neurones dans la couche de sortie correspondant aux différentes classes.
- La fonction de transfert $tansig$ (Tangente sigmoïde) pour la couche cachée, et la fonction $purelin$ (Linéaire) pour la couche de sortie.
- L'apprentissage se fait par l'algorithme de la rétropropagation du gradient basé sur l'Algorithme de Levenberg Marquardt (`Trainlm`).
- Fonction de performance : mse (Mean square error- Erreur moyenne quadratique).
- Les valeurs des paramètres suivants: nombre des neurones dans la couche cachée et le nombre maximal d'itérations, l'erreur minimale, et le taux d'apprentissage ne sont pas définis, car l'objectif de cette partie du travail est d'étudier leurs effets sur les performances du RNA .

Le critère qui nous permet d'évaluer les performances du réseau est le pourcentage de classification (taux de succès) T .

$$T = \frac{N_{Bc}}{N} \times 100$$

Avec

N_{BC} : Nombre d'observations bien classées.

N : Nombre total d'observations.

4.3.3- Effet du nombre de neurones dans la couche cachée :

Afin de choisir l'architecture pour la quelle, on obtient un meilleur pourcentage d'observations bien prédites en test, plusieurs architectures sont étudiées; en variant le nombre des neurones qu'elle contienne de 1 à 20 neurones. Le nombre des neurones dans la couche d'entrée est fixé à 26 et celui de la couche cachée est fixé à 2. Le réseau est entraîné en utilisant la configuration citée précédemment. Nous faisons plusieurs essais (Total de 10 essais) pour chaque nombre de neurones. Puis, nous calculons la moyenne et l'écart type du taux de classification correspondant à chaque nombre de neurones. Les résultats obtenus sont présentés dans le **tableau 4-7**:

Nombre des neurones cachés	Pourcentage des bien classés (Apprentissage)	Ecart- type Apprentissage	Pourcentage des bien classés (Test)	Ecart- type Test
1	24	12.43	24	12.43
2	56	12.64	56	12.64
3	88	10.32	81.2	10.55
4	100	0	99.8	0.63
5	100	0	99.6	1.26
6	100	0	99.8	0.63
7	100	0	100	0
8	100	0	100	0
9	100	0	100	0
10	100	0	100	0
15	100	0	100	0
20	100	0	100	0

Tableau 4.7- *Effet du nombre de neurones dans la couche cachée sur le pourcentage d'observations bien classées.*

Discussion:

L'allure du pourcentage des observations bien prédites test/ Apprentissage en fonction du nombre de neurones dans la couche cachée est représentée sur la figure 4.8. On constate que la performance de l'apprentissage augmente vis-à-vis du nombre de neurones dans la couche cachée, jusqu'à la valeur maximale (100%), qui correspond à un nombre de neurones égal à 4. Mais, elle reste constante après cette valeur. Généralement, pour juger les performances du RNAs en généralisation, on se réfère à la courbe des biens classés du test. Le nombre de neurones dans la couche cachée est ainsi fixé par cette courbe où nous avons remarqué que la performance augmente jusqu'à atteindre la valeur maximale (100 %), qui correspond à un nombre de 7 neurones.

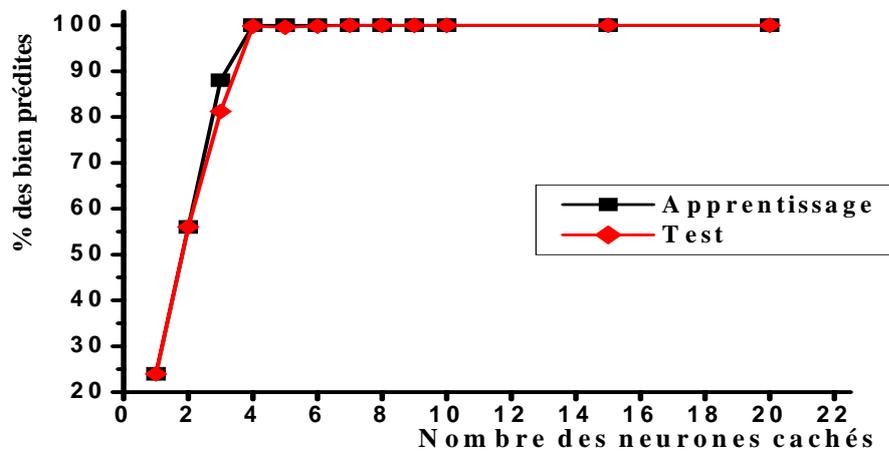


Fig.4.8- Effet du nombre des neurones dans la couche cachée sur le pourcentage d'observations bien classées.

Nous avons donc choisi, un nombre optimal de neurones dans la couche cachée de 5 pour la suite des essais.

4.3.4- Choix des valeurs optimales des critères d'arrêt (erreur minimale et nombres des itérations maximales):

Dans cette section, nous étudions l'effet des différentes valeurs du nombre des itérations maximales et de l'erreur minimale ; en variant le premier paramètre dans un intervalle de $[1,100]$ et le deuxième paramètre dans un intervalle de $[10^{-1},10^{-7}]$. Nous faisons plusieurs essais (total de 10 essais) pour chaque paramètre. Puis, nous calculons la moyenne et l'écart type du taux de classification (Apprentissage et test) correspondant à chaque nombre de neurone.

a- Effet du nombre d'itérations maximales:

Les résultats de l'effet du nombre d'itérations maximales sont représentés sur le *tableau 4.8*:

Nombre d'itérations maximales	Pourcentage des bien classés (Apprentissage)	Ecart- type Apprentissage	Pourcentage des bien classés (Test)	Ecart- type Test
1	39.8	17.93	39.8	13.44
2	64.77	12	62	14.66
5	96.11	12.29	95	14.42
10	100	0	100	0
20	100	0	100	0
50	100	0	100	0
100	100	0	100	0

Tableau 4.8- Effet du nombre d'itérations maximales sur le pourcentage d'observations bien cassées.

Discussion:

Sur la figure 4.9, nous représentons la variation des pourcentages bien prédites apprentissage et test, afin de montrer l'effet du choix de la valeur du nombre d'itérations maximales. On remarque que la performance de l'apprentissage augmente en fonction du nombre d'itérations maximales, jusqu'à la valeur maximale (100%), qui correspond à un nombre 10 itérations. Après ce point, elle reste constante. Pour mesurer les performances du RNAs en généralisation, on se réfère à la courbe des biens classés du test. Le nombre d'itérations est ainsi fixé par cette courbe où nous avons remarqué que la performance augmente jusqu'à atteindre la valeur maximale (100 %), qui correspond à un nombre de 10 itérations.

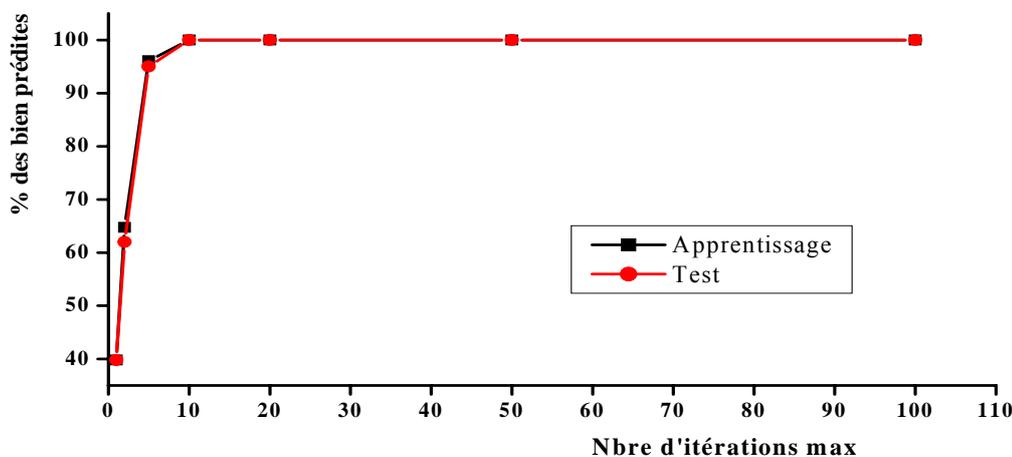


Fig4.9- Effet du nombre d'itérations maximales sur le pourcentage d'observations bien cassées.

Alors, nous avons choisi, un nombre optimal d'itérations maximales de 10 itérations pour le reste de l'application.

b- Effet de l'erreur minimale:

Les résultats de l'effet du choix de la valeur de l'erreur minimale (mse) sont représentés sur le tableau 4.9:

Erreur minimale (mse)	Pourcentage des bien classés (Apprentissage)	Ecart- type Apprentissage	Pourcentage des bien classés (Test)	Ecart- type Test
10^{-7}	100	0	100	0
10^{-6}	100	0	100	0
10^{-5}	100	0	100	0
10^{-4}	100	0	100	0
10^{-3}	100	0	99.8	0.63
10^{-2}	98.88	0.35	99.6	0.84
10^{-1}	81.44	17.67	78.6	16.76

Tableau 4.9- Effet de l'erreur minimale

- Discussion:

L'effet de l'erreur minimale est schématisé sur la figure 4.10, qui représente l'allure du taux des observations bien classées d'apprentissage et de test. En effet, l'analyse des courbes montre que la performance en apprentissage et en test est maximale (100%), pour une valeur de l'erreur égal 10^{-4} .

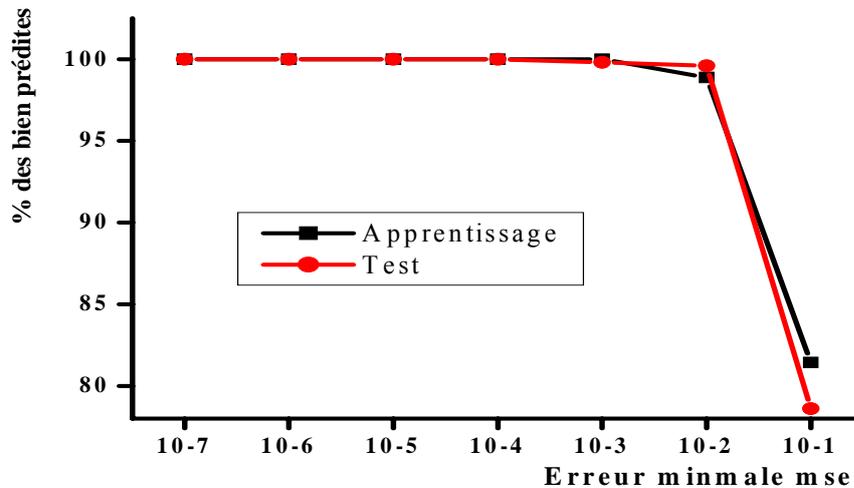


Fig4.10- Effet de l'erreur minimale sur le pourcentage d'observations bien cassées.

Puisque, la mesure de la capacité en généralisation d'un RNA, se base sur la performance de test. Nous avons donc choisi, à partir de la courbe performance de test, une valeur optimale de l'erreur minimale mse de 10^{-4} pour la suite du travail.

3.3.5- Effet du taux d'apprentissage:

Afin d'étudier l'effet de la valeur du taux d'apprentissage par la recherche de sa valeur optimale qui correspond à la meilleure performance du réseau de neurones. Nous avons fait une série d'essais (total de 10 essais), en variant la valeur du taux d'apprentissage η dans un intervalle de $[0.1, 0.9]$, puis nous calculons la moyenne et l'écart type. Les résultats obtenus sont représentés sur le *tableau 4.10*.

Taux d'apprentissage	Pourcentage des bien classés (Apprentissage)	Ecart- type Apprentissage	Pourcentage des bien classés (Test)	Ecart- type Test
0.1	99.88	0.35	98.6	3.13
0.2	99.88	0.35	99.3	2.52
0.3	100	0	99.4	0.96
0.4	100	0	99	3.16
0.5	100	0	100	0
0.6	100	0	99.8	0.63
0.7	99.3	2.10	99.3	2.10
0.8	100	0	99.8	0.63
0.9	99.44	1.75	97.4	6.18

Tableau 4.10- *Effet du taux d'apprentissage.*

- Discussion:

L'allure du pourcentage des observations, bien prédites test/ Apprentissage variant en fonction de la valeur du taux d'apprentissage (η), est schématisée sur la figure **4.11**. On remarque que la performance de l'apprentissage présente généralement une droite avec l'augmentation de η . Pour juger les performances des RNAs en généralisation, on se réfère à la courbe des biens classés du test. Le taux d'apprentissage permettant d'avoir une meilleure performance (100 %), égal à 0.5.

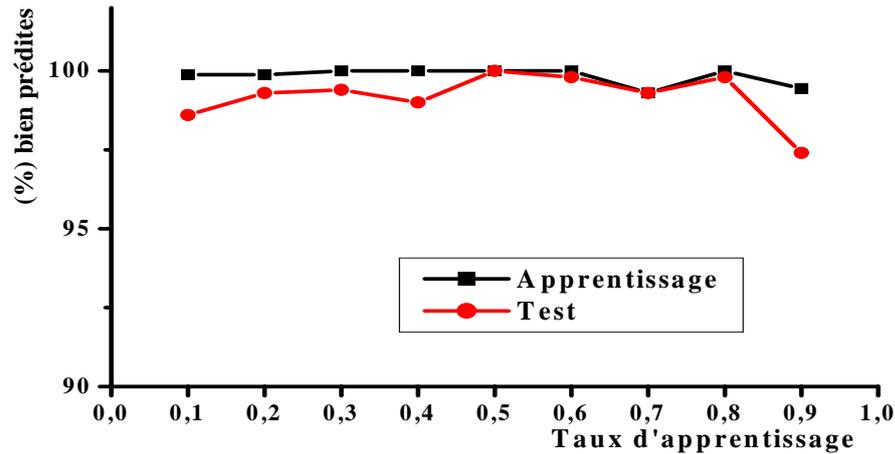


Fig.4.11- Effet du taux d'apprentissage sur le taux d'observations bien classées.

Nous avons donc choisi, à partir de la courbe performance de test, la valeur du taux d'apprentissage (η) de 0.5.

4.4- Conclusion:

Dans ce chapitre, nous avons présenté les différents résultats obtenus lors de l'utilisation du réseau de neurones perceptron multicouches (PMC) pour automatiser le diagnostic des défauts d'engrenages et de roulements.

Nous avons aussi étudié l'effet du choix des différents paramètres (configuration) du réseau PMC sur sa performance de classification. Dans un premier temps, nous avons réalisé un prétraitement des signaux vibratoires prélevés des deux bancs d'essais (CETIM et Bearing Data Center), pour extraire les indicateurs de surveillance les plus pertinents à l'état de santé des deux dispositifs expérimentaux. Puis, nous avons construit la base de données utilisée pour l'apprentissage et le test du PMC. Nous avons déterminé, aussi, la configuration optimale du PMC appropriée à chacun des deux bancs d'essai étudiés.

Lors de nos essais, nous avons remarqué que l'augmentation du nombre des neurones dans la couche cachée, et par conséquent l'augmentation de la taille de la couche cachée, rend l'apprentissage très lent et influe d'une manière négative sur la convergence de l'algorithme de la rétropropagation du gradient. En ce qui concerne les critères d'arrêt; la valeur de l'erreur minimale ne doit pas être très petite et le nombre d'itérations maximales, aussi, ne doit pas être très grand. Car, si ces deux valeurs sont mal choisies, on risque de la saturation du réseau et le processus d'apprentissage ne

s'arrête jamais. Pour la valeur du taux d'apprentissage, elle influe directement sur l'ajustement des poids synaptiques.

En résumé, les valeurs optimales des paramètres du PMC sont:

1. Pour le banc CETIM (engrenage) à 2 classes
 - Nombre d'itérations maximales égal 200.
 - Nombre de neurones dans la couche cachée égal 5.
 - Taux d'apprentissage égal 0.8.
 - Erreur minimale égal 10^{-3} .
2. Pour le banc d'essai Bearing data center (roulement) à 5 classes
 - Nombre d'itérations maximales égal 10.
 - Nombre de neurones dans la couche cachée égal 7.
 - Taux d'apprentissage égal 0.5.
 - Erreur minimale égal 10^{-4} .

CONCLUSION

GENERALE

CONCLUSION GENERALE

Le travail effectué dans le cadre de cette thèse s'inscrit dans le domaine de la surveillance et le diagnostic automatisés des systèmes mécaniques par analyse vibratoire.

Après avoir présenté un aperçu théorique, dans lequel nous avons montré les défauts courants dans les machines tournantes et les différentes méthodes de diagnostic par analyse vibratoire. Deux approches sont citées, en précisant leurs avantages et leurs inconvénients : l'approche classique basé sur le jugement (décision) de l'homme sur l'état de santé des machines par l'analyse des signaux vibratoire, en utilisant les différentes méthodes du traitement de signal. La deuxième approche dite moderne; est basée sur l'automatisation de la décision, en utilisant les différents algorithmes et règles de décision et surtout celles de l'intelligence artificielle. Nous avons tout d'abord étudié l'état d'art de l'application des réseaux de neurones artificiels (RNAs) au diagnostic des défauts des machines tournantes. On a présenté la synthèse de quelques travaux réalisés sur ce sujet, dont le but était de savoir la méthodologie de l'utilisation des réseaux de neurones pour automatiser le diagnostic vibratoire. Le réseau de neurone le plus utilisé en surveillance est le perceptron multicouches (PMC). Dans la suite du travail, nous avons également abordé, de manière approfondie la théorie des réseaux de neurones et particulièrement le réseau PMC. Aussi, on a présenté les étapes nécessaires pour implémenter le PMC sous Matlab et la démarche adoptée pour l'utiliser comme un outil de diagnostic. Nous avons expliqué en détail, le choix des différents paramètres pertinent à la structure et au fonctionnement du PMC. Ensuite, nous avons présenté les bancs d'essai étudiés : banc d'essai d'engrenages (CETIM), et le banc d'essai de roulement (Data Bearing Center), et les signaux vibratoires correspondants.

Enfin, nous avons présenté et interprété les différents résultats expérimentaux en jouant sur la diversité des paramètres du réseau et la structure de la base de données, et en utilisant les programmes que nous avons développé. Nous nous sommes

intéressés à l'effet des paramètres suivants : critères d'arrêt, nombres de neurones dans la couche cachée et le taux d'apprentissage. Ces paramètres ont une influence directe sur le processus d'apprentissage du réseau. L'étude de l'effet de ces paramètres sur la performance en classification du réseau PMC, nous a conduit à trouver la configuration optimale appropriée à chaque banc d'essai. Aussi, On a montré que pour réussir un tel diagnostic vibratoire basé sur la méthode des réseaux de neurones artificiels, il faut :

- Bien choisir et traiter les descripteurs (indicateurs de surveillance) formant la base de données utilisée pour entraîner et tester le PMC.
- Bien choisir les valeurs des paramètres et les fonctions du réseau PMC.

Par ailleurs, un des problèmes encore ouverts est celui de l'évaluation des performances de généralisation des réseaux de neurones utilisés en diagnostic des systèmes mécanique; qui nécessite toujours une grande base de données. Ce qui n'est pas disponible pratiquement. En effet, pour avoir suffisamment de données, il faut augmenter le nombre des essais (expériences). Et cela, nécessite la détérioration d'un grand nombre d'éléments (roulements, engrenage, pompes, etc.). Par conséquent, la facture des essais devient trop élevée, ce qui n'est pas à la portée de tous les intéressés par ce domaine. La modélisation des dispositifs et des systèmes mécaniques, malgré sa difficulté, peut résoudre ce problème et fournir les données nécessaires pour réussir un diagnostic basé sur les réseaux de neurones artificiels.

***REFERENCES
BIBLIOGRAPHIQUES***

REFERENCES BIBLIOGRAPHIQUES

[ALATT94], Alattas Mohamed, « *Maintenance des machines tournantes : signature de défauts d'engrenages droit et hélicoïdaux* », Thèse de l'INSA de Lyon, 1994, pp.198.

[BAGHL99], Lotfi Baghli, « *Contribution à la commande de la machine asynchrone, utilisant de la logique flou, des réseaux de neurones et des algorithmes génétiques* », Thèse de doctorat, U- Henri Poincaré-Nancy-1, 1999, pp.212.

[BENOU99], Benouaghram Ramzi, « *Contribution au diagnostic des machines tournantes complexes. Application à un laminoir* », Thèse de l'UTC de Compiègne, Avril 1999, pp.128.

[BISH95], C. M. Bishop, « *Neural networks for pattern recognition* », Oxford Univ. Press, 1995, pp.498.

[BEOK93], Mathew J. Boel, Jacob L. Cybulski and Andre S. Szczepanik, « *Embedding Neural Networks in On-line Monitoring Applications* », ANZIIS'93.

[BOUL03] Boulanger A., Pachaud C., « *Analyse vibratoire en maintenance. Surveillance et diagnostic des machines* », 2ème édition, Edition Dunod, Paris 2003, pp. 365.

[BOUI00], Bouillaut Laurent, « *Approches cyclostationnaire et non-lineaire pour l'analyse vibratoire de machines tournantes : Aspect théoriques et applications au diagnostic* », Thèse de l'UTC de Compiègne, Novembre 2000, pp.183.

[BOUZ03], B. Bouzouane, A. Miloudi, N. Hamzaoui & A. Benchaala, « *Détection de défauts de machines tournantes par la méthode des ondelettes* », 16ème Congrès Français de Mécanique Nice, 1-5 septembre 2003, pp.1-6.

[BREN02], Breneur Claire « *Éléments de maintenance préventive de machines tournantes dans le cas de défauts combinés d'engrenage et de roulements* », Thèse de l'INSA de Lyon, Décembre 2002, pp.120.

[BROE06], Piet M.T. Broersen, « *Automatic Autocorrelation and Spectral Analysis* », Springer-Verlag London Limited 2006, pp.298.

[CARRE90], D. Carreau, « *Surveillance des roulements par analyse des vibrations* ». CETIM-informations no1S, avril 1990.

[DUDA03], Duda R.O., Hart P.E., Stork D, « *Pattern Classification* ».2ème édition., Ricoh Research Center California 1997, pp. 738.

[ELBADA99a], M. EL BADAOU, F. GUILLET, J. DANIÈRE, « Surveillance des systèmes complexes à engrenages par l'analyse cepstrale synchrone », *Traitement du Signal* 1999 - Volume 16 - n°5, pp.371-381.

[ELBADA99b], Mohamed EL BADAOU, « *Contribution au Diagnostic Vibratoire des Réducteurs Complexes à Engrenages par l'Analyse Cepstrale* ». Thèse de doctorat à L'UNIVERSITE JEAN-MONNET, juillet 1999, pp.141.

[FEDA05], FEDALA Semchedine, « *Le diagnostic vibratoire automatisé : comparaison des méthodes d'extraction et de sélection du vecteur forme* », Mémoire de magister DOMP-UFAS, Sétif, 2005, pp.142.

[FELK95], FELKAOUI Ahmed, « Contribution à l'identification des signaux et des systèmes dynamiques linéaires », Thèse de magister, université de Sétif 1995.

[FELK07], FELKAOUI Ahmed, « Diagnostic vibratoire au service de la maintenance : de l'acquisition à la décision », Thèse de doctorat DOMP-UFAS, 2007, pp.168.

[HAYK01], Simon HAYKIN, « *Neural Networks - A Comprehensive Foundation* ». Second Edition, Pearson Prentice Hall, India 2001, pp.823.

[HOWA00], Howard Demuth , Mark Beale, « *Neural Network Toolbox For Use with MATLAB, User's Guide Version 4* », Copyright 1992 - 2000 by the MathWorks, Inc.

[GUPT03], Madan M.Gupta, Liang Jin, and Noriyasu Homma, « *Static and Dynamic Neural Networks from Fundamentals to Advanced Theory* », John Wiley & sons, inc. publication 2003, pp.751.

[JACK00], L. B. Jack, A. K. Nandi, « *Feature selection for ANNs using genetic algorithms in detection of bearing faults* », IEE proceeding vision, image and signal processing, 147(3), 2000, pp.205-212.

[KECM01], Vojislav Kecman, « *Learning and soft computing : support vector machines, neural networks, and fuzzy logic models* », The MIT Press, Cambridge, The MIT Press Cambridge, London, 2001, pp.568.

[KIRKH97], Chris KIRKHAM, Tom Harris, « *A hybrid neural network system for generic bearing fault detection* », Proceedings of COMADEM 97.

[LEPI 01], Ülo LEPIK, « *Application of wavelet transform techniques to vibration studies* », Proc. Estonian Acad. Sci. Phys. Math., 2001, pp155–168.

[LEBO00], Lebold M., McClintic K.; Campbell R., Byington C., Maynard K., « *Review of Vibration Analysis Methods for Gearbox Diagnostics and Prognostics* », Proceedings of the

54th Meeting of the Society for Machinery Failure Prevention Technology, Virginia, May 1-4, 2000, pp. 623-634.

[MAcCR96a], A.C. McCormick, A. K. Nandi, « *Rotating Machine Condition Classification Using Artificial Neural Networks* », Proceedings of COMADEM '96, University of Sheffield, 16th-18th July 1996.

[MAcCR96b], A.C. McCormick, A. K. Nandi, « *A comparison Of artificial neural networks and other statistical methods for rotating machine condition classification* », IEE colloquium on Modeling and Signal Processing for Fault Detection, University of Leicester, 18th September 1996.

[MAHF00], J. MAHFOUDH, C. BRENEUR, F. CALDERONE et D. PLAY, « *Assessment of neural networks for detection and identification of gearbox faults* », ESDA 2000, the Centennial Interdisciplinary Conference on Research, Systems Integration, Design & analysis, Montreux / SWIZERLAND, 10-13 July 2000.

[Marq01] J. P. Marques de Sa, « *Pattern Recognition Concepts, Methods and Applications* », Springer 2001, pp.328.

[MICH94], D. Michie, D.J. Spiegelhalter, C.C. Taylor « *Machine Learning, Neural and Statistical Classification* », February 17, 1994, pp. 298.

[MILI06], X. MILISEN, V. LENAERTS, « *Comparaison de différentes méthodes de détection de défauts dans les roulements* », Revue Scientifique des ISILF n°20, 2006, pp.157-171.

[NORF06], Derek Norfield, « *Practical Balancing of Rotating Machinery* ». Elsevier2006. pp.217.

[OEHLM95], Oehlmann H, Brie D., Begotto V., Tomczak M., « *Analyse temps fréquence de l'écaillage d'engrenage de boîte de vitesses* », 2ème conf. Inter. Méthode de surveillance, Senlis, France ,10-12 October 1995 vol. 1, pp. 243-253.

[PADO99], Linilson R. Padovese, « *Using acoustical noise for fault classification in gearbox* », 15th Brazilian congress of mechanical Engineering (COBEM99), 22-26 November 1999.

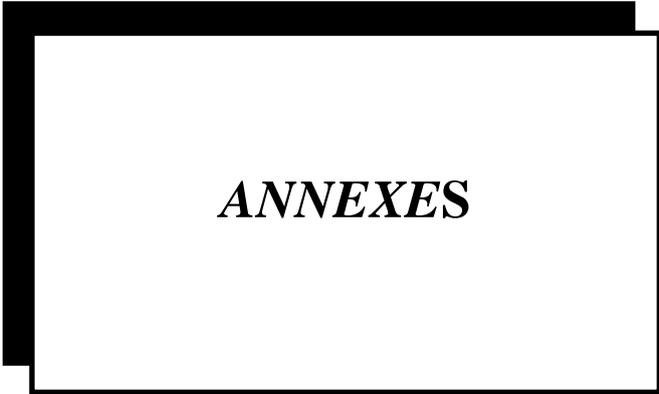
[PARIZ04], Marc PARIZEAU, « *Réseaux de neurones* », Université LAVAL automne 2004, pp.127.

[PAYA97], b.a.PAYA, i.i ESAT et m.n. BADI, « *Artificial neural network based fault diagnostics of rotating machinery using wavelet transform as a preprocessor* », Mechanical Systems and Signal Processing 1997, 11(5), Academic Press , pp. 751-765.

[PING00], Ping Chen, « *Bearing Condition Monitoring and Fault Diagnosis* », master of science thesis, the university of Calgary, Canada, 2000, pp.132.

- [RACO06], Daniel RACOCEANU, « *Contribution à la surveillance des Systèmes de Production en utilisant les Techniques de l'Intelligence Artificielle* », Habilitation à diriger des recherches, université de Franche-Comté de besançon. 2006, pp.240.
- [ROHE 90], N.M.ROEHI, C.E.PEDREIRA, M.SFIALHO, « *Incipient Fault Detection in Rotating Machinery by Using Artificial Neural Networks* », Proceedings on fault detection supervision and safety for technical processes, Finland, 1990.
- [ROUG94], Rouget D., « *De l'analyse d'avarie à la conception des montages de roulements* ». Revue Française de Mécanique, 1994, pp.181-186.
- [SAMA04], B. Samanta, « *Gear fault detection using artificial neural networks and support vector machines with genetic algorithms* », Mechanical Systems and Signal Processing 18 (2004), pp.625–644.
- [SHIGE05], Shigeo Abe, « *Support Vector Machines for Pattern Classification* », Springer-Verlag London Limited 2005, pp.349.
- [SIDAH 90], Sidahmed M., « *Détection précoce des défauts dans les engrenages par analyse vibratoire* », 4^{ème} journée d'étude des transmissions mécaniques, juin ,1990.
- [SORS91], Timo Sorsa, Heikki N.Koivo, Hannu koivisto, « *Neural networks in fault diagnosis* », iee transactions on systems, man and cybernetics vol.21 no. 4, July /august 1991, p815-825.
- [SUN98], Q. SUN, F.XI, and G.Krishnappa, « *Signature analysis of rolling element defects* », Proceeding of CSME Forum,Toronto1998, pp.423-429.
- [STAR05], Stark Hans-Georg, « *Wavelets and Signal processing* », Springer, Berlin, 2005, pp.150.
- [THEOD03], Sergios Theodoridis, konstantinos koutroumbas, « *Pattern recognition* », second edition, Copyright 2003, Elsevier (USA), pp.710.
- [THOM03], Marc Thomas, « *Les vibrations des paliers fluides et d'engrenages* », cours de Techniques de maintenance industrielle, ETS, École de technologie supérieure, Québec, Canada, Novembre 2002.
- [VAPN95], V. Vapnik, « *The Nature of Statistical Learning Theory* », Springer Verlag 1995.
- [YANG00], Da-Ming Yang et James Penman, « *Intelligent detection of induction motor bearing faults using current and vibration monitoring* ». Proceeding of COMADEM 2000, 3-8 December 2000 Houston Texas USA, pp.461-470.
- [ZEMO03], Zemouri, R., « *Contribution à la surveillance des systèmes de production à l'aide des réseaux de neurones dynamiques : Application à la e-maintenance* », Thèse de Doctorat, Université de Franche Comté, Besançon.2003, pp.277.

[ZWIN95], Zwingelstein G., « *Diagnostic des défaillances. Théorie et pratique pour les systèmes industriels* », Traité des Nouvelles Technologies - Série Diagnostic et Maintenance. Editions Hermès, Paris 1995, pp.600.



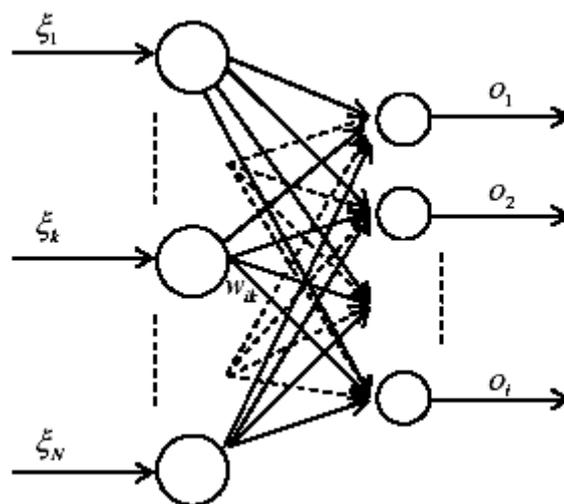
ANNEXES

ANNEXE I

QUELQUES TYPES USUELS DE RÉSEAUX DE NEURONES

1. Le perceptron simple

C'est un des premiers réseaux de neurones, conçus en 1958 par Rosenblatt. Il est linéaire et monocouche. Il est inspiré du système visuel. La première couche (d'entrée) représente la rétine. Les neurones de la couche suivante (unique, d'où le qualificatif de monocouche) sont les cellules d'association, et la couche finale les cellules de décision. Les sorties des neurones ne peuvent prendre que deux états (-1 et 1 ou 0 et 1).



FigA1.1- Perceptron Simple, modèle de Rosenblatt.

$$O_i = f(a_i) = f\left(\sum_{k=1}^{k=N} w_{ik} \xi_k\right)$$

Où w_{ik} est le poids de la connexion qui relie l'unité k à l'unité i , a_i est l'activation de l'unité i , f est la fonction d'activation des unités (Figure 16). Cette fonction d'activation est du type fonction à seuil avec l'expression suivante :

$$f(x) = \begin{cases} +1 & \text{si } x \geq \theta \\ -1 & \text{si } x < \theta \end{cases}$$

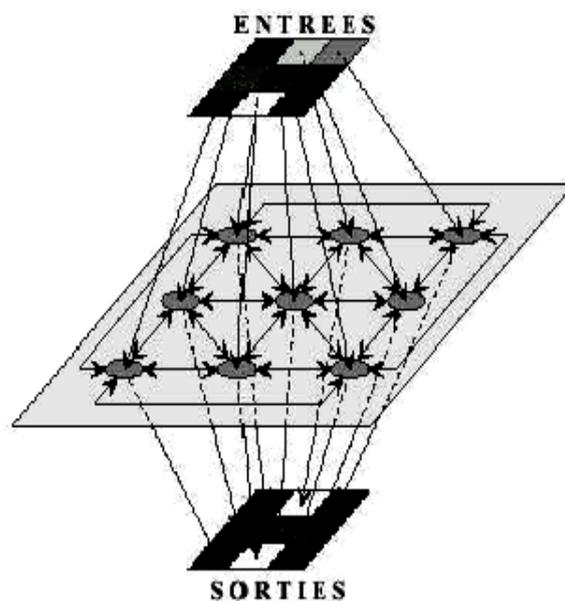
Seuls les poids des liaisons entre la couche d'association et la couche finale peuvent être modifiés. La règle de modification des poids utilisée est la règle de Widrow-Hoff : si la sortie du réseau (donc celle d'une cellule de décision) est égale à la sortie désirée, le poids de la connexion entre ce neurone et le neurone d'association qui lui est

connecté n'est pas modifié. Dans le cas contraire le poids est modifié proportionnellement à la différence entre la sortie obtenue et la sortie désirée.

En 1969, Papert et Minsky ont démontré les limites du perceptron classique, incapable, par exemple de simuler la fonction ou exclusif (xor).

2. Les réseaux de Hopfield

Il s'agit d'un réseau constitué de neurones à deux états (-1 et 1, ou 0 et 1), dont la loi d'apprentissage est la règle de Hebb (1949), qui veut qu'une synapse améliore son activité si et seulement si l'activité de ses deux neurones est corrélée (c'est-à-dire que le poids d'une connexion entre deux neurones augmente quand les deux neurones sont activés au même temps).



FigA1. 2- Le modèle de Hopfield utilisé comme mémoire adressable par le contenu

3. Les réseaux de Kohonen

Contrairement aux réseaux de Hopfield où les neurones sont modélisés de la façon la plus simple possible, on recherche ici un modèle de neurone plus proche de la réalité. Ces réseaux sont inspirés des observations biologiques du fonctionnement des systèmes nerveux de perception des mammifères.

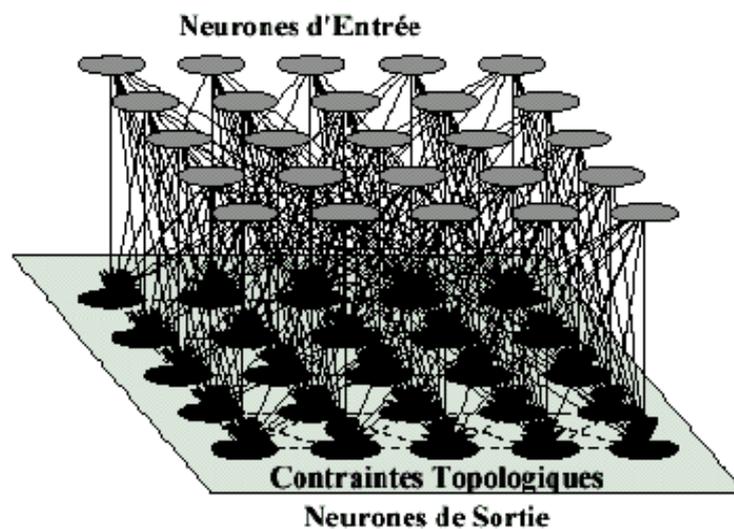
Une loi de Hebb modifiée (tenant compte de l'oubli) est utilisée pour l'apprentissage. La connexion est renforcée dans le cas où les neurones reliés ont une activité simultanée, et diminuée dans le cas contraire (alors qu'il ne se passait précédemment rien dans ce cas). Ceci se résume par la formule :

$$dw / dt = k S e - B(S) w$$

Où w est le poids associé à une certaine connexion, e la valeur que le neurone reçoit en entrée par cette connexion, S la valeur qu'il renvoie en sortie (toujours positive), $B(S)$ la fonction d'oubli et k une constante positive.

Une loi d'interaction latérale (Observée biologiquement) est aussi modélisée. Les neurones très proches (physiquement) interagissent positivement (le poids des connexions est augmenté autour d'une certaine zone quand une synapse est activée), négativement pour les neurones un peu plus loin, et pas du tout pour les neurones éloignés. Ceci crée un "amas" de neurones activés et contribue à spécialiser certains neurones : pour une entrée donnée, une sortie particulière sera activée alors que les autres resteront inertes. On utilise aussi parfois des lois de concurrence entre les neurones (création et destruction de neurones selon certains critères).

Ceci permet de résoudre certains problèmes, tels le problème du voyageur de commerce (comment relier n villes par le chemin le plus court).

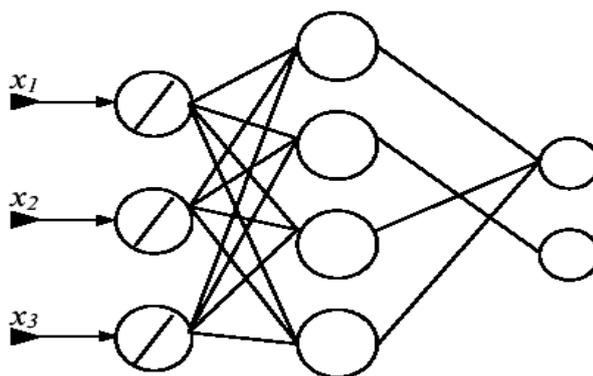


FigA1.3- *Connectivité dans les cartes topologiques de Kohonen, chaque cellule de la couche supérieure est reliée à l'ensemble des cellules de la carte inférieure*

Les réseaux de Kohonen ont des applications dans la classification, le traitement de l'image, l'aide à la décision et l'optimisation.

4. Réseaux de neurones à fonction de base radiale (RBF)

Les réseaux de neurones à fonction de base radiale (Radial Basis Funiculites) sont des réseaux de neurones à une seule couche cachée dont les fonctions d'activation sont des fonctions à base radiale, le plus souvent des gaussiennes. La fonction d'activation du neurone de la couche de sortie est l'identité. Les entrées sont directement connectées aux neurones de la couche cachée.



FigA1.4- Réseau à fonctions de base radiales.

ANNEXE 2

FONCTIONS DE TRANSFERT

Les fonctions d'activations, les plus courantes et les plus citées dans la littérature, ainsi leurs noms sous *Matlab*, sont énumérées dans la figure ci-dessous, Avec : $a=f(n)$

Nom de la fonction	Relation d'entrée/sortie	Icône	Nom Matlab
seuil	$a = 0$ si $n < 0$ $a = 1$ si $n \geq 0$		hardlim
seuil symétrique	$a = -1$ si $n < 0$ $a = 1$ si $n \geq 0$		hardlims
linéaire	$a = n$		purelin
linéaire saturée	$a = 0$ si $n < 0$ $a = n$ si $0 \leq n \leq 1$ $a = 1$ si $n > 1$		satlin
linéaire saturée symétrique	$a = -1$ si $n < -1$ $a = n$ si $-1 \leq n \leq 1$ $a = 1$ si $n > 1$		satlins
linéaire positive	$a = 0$ si $n < 0$ $a = n$ si $n \geq 0$		poslin
sigmoïde	$a = \frac{1}{1+\exp^{-n}}$		logsig
tangente hyperbolique	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$		tansig
compétitive	$a = 1$ si n maximum $a = 0$ autrement		compet

FigA2.1- Les fonctions d'activation les plus utilisées.

Résumé:

L'objectif de ce mémoire est d'appliquer les réseaux de neurones artificiels à l'automatisation du diagnostic des machines tournantes par analyse des signaux vibratoires. Cette technique est basée sur la classification par reconnaissance de formes, elle a pour but d'augmenter la précision (l'exactitude) et de réduire les erreurs causées par un jugement subjectif de l'homme.

L'efficacité de cette procédure repose sur les quatre étapes suivantes, à savoir : l'acquisition du signal vibratoire, le pré-traitement des données, l'extraction des caractéristiques et enfin la reconnaissance de ces caractéristiques par les réseaux de neurones. Une attention particulière est donnée au choix de l'architecture du réseau, des algorithmes d'apprentissage et des méthodes d'évaluation des performances.

Mots clés:

Machines tournantes, Vibrations mécaniques, Réseaux de neurones artificiels, Diagnostic des défauts, Traitement de signal.

Abstract :

The objective of this thesis is to apply the artificial neural networks to the automated diagnostic of the rotating machines by analyzing of the vibratory signals. This technique is based on classification by pattern recognition; the purpose is to increase the precision (exactitude) and to reduce the errors caused by a man subjective judgment.

The effectiveness of this procedure dependence on the four following stages, namely: vibratory signal acquisition, data preprocessing, characteristics extraction and finally the recognition of these characteristics by the neural networks. A particular attention is given to the choice of network architecture, the training algorithm, and the evaluation method of network performances.

Keywords:

Rotating machines, Mechanical vibrations, Artificial Neural Networks, Fault Diagnosis, signal processing.

المخلص:

إن هدف هذه المذكرة هو استخدام الشبكات العصبية الاصطناعية في التشخيص الآلي لأعطال الآلات الدوارة، بتحليل الإشارات الاهتزازية. تعتمد هذه التقنية على التصنيف بواسطة تمييز الأشكال. وهدفها زيادة الدقة و التقليل من الأخطاء التي يسببها قرار الإنسان.

فعالية هذا الإجراء تعتمد على المراحل الأربعة التالية: اقتناء الإشارة الاهتزازية، تحضير المعطيات، استخراج الميزات و أخيرا الانتباه الدقيق لاختيار هندسة الشبكات العصبية، خوارزمية التدريب و طرق تقييم الأداء.

الكلمات المفتاحية:

الآلات الدوارة، الاهتزازات الميكانيكية، الشبكات العصبية الاصطناعية، تشخيص الأعطاب، تحليل الإشارة.