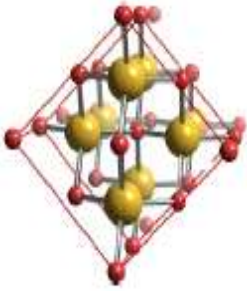


REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHESCIENTIFIQUE
UNIVERSITE ABOU-BAKR BELKAID DE TLEMCEN



FACULTE DES SCIENCES

Département de Physique



Laboratoire d'Etude et Prédiction de Matériau
Unité de recherches Matériaux et Energies Renouvelables

MEMOIRE DE MASTER

EN PHYSIQUE de la MATIERE CONDENSEE et des SEMICONDUCTEURS

Présentée par :

M^{elle} Kadous Djamila

THEME

Utilisation des réseaux de neurones comme outil du
datamining :Génération de modèle comportemental d'un
processusphysique à partir de données

Soutenu le :28/06/2012 devant le jury :

Pr. T. BOUCHAOUR

Président

Université de Tlemcen

Dr. A. SENOUDI

Examinatrice

Université de Tlemcen

Pr. A. RAHMOUN

Encadreur

Université de Sidi Bel Abbes

Année Universitaire : 2011/2012

Dédicaces

*Avec la plus grande joie et cœur aimant que je dédie ce
travail :*

A la grâce de dieu qui m'a guidé

*A mes défunts parents, que j'aurais aimé qu'ils
assistent à ce jour*

A mes sœurs Nacera, Fatiha, Mériem

A mes nièces Sabrina et Ryme

A toute ma famille

A mes amis qui me sont chères.

Remerciements

Je remercie ALLAH le Tout-puissant de m'avoir donné le courage, la volonté et la patience de mener à terme ce présent travail.

*Ce travail a été réalisé dans le laboratoire de recherche d'Etudes et Prédiction des Matériaux (LEPM) affilié à l'Unité de Recherche des Matériaux et des Energies Renouvelables (URMER) de la faculté des sciences, sous la direction de Monsieur **G. MERAD**, Professeur à l'Université de Tlemcen. Je lui exprime toutes mes reconnaissances de m'avoir accepté dans son laboratoire et sa persévérance de continuer.*

*J'exprime mes sincères remerciements à Monsieur **T. BOUCHAOUR**, Professeur à l'Université de Tlemcen, d'avoir accepté de présider le Jury de cette thèse.*

*Je tiens vivement à exprimer ma gratitude et mes remerciements à Monsieur **A. RAHMOUN**, Professeur à l'Université de Sidi Bel-Abbes, pour son aide précieuse, ses encouragements et ses conseils tout le long de ce travail.*

*Mes remerciements les plus distingués pour Melle **A. SENOUDI**, Docteur d'état à l'université de Tlemcen, pour ses conseils tout le long du master M1 et d'être membre de jury.*

Enfin, j'adresse mes vives reconnaissances à tous les membres de l'unité de recherches (URMER) sans exception, sans oublier les étudiants de Master de la matière condensée ainsi que mes amis fonctionnaires de la direction des ressources en eau de la wilaya de Tlemcen qui m'ont encouragé et soutenu jusqu'au bout.

Sommaire

Introduction générale	01
Chapitre I : Data mining	
I-1 : Introduction	05
I-2 : Qu'est-ce que le datamining	05
I-3 : Etapes d'un projet de datamining	07
I-4 : les Tâches du datamining	08
I-4-1 : La classification	08
I-4-2 : Clustering	08
I-4-3 : Règles d'association	08
I-4-4 : Prédiction	09
I-5 : Les techniques du datamining	09
I-5-1 : Les techniques descriptives	09
I-5-1-1: Analyse en composantes principales ACP	10
I-5-1-2: Les k-means	10
I-5-1-3 : La classification neuronale	11
I-5-1-4 : Domaines d'application	13
I-5-2 : Les techniques prédictives	14
I-5-2-1 : Les arbres de décision	15
I-5-2-2 : La classification supervisée	16
I-6 : Conclusion	18
CHAPITRE II : Les réseaux de neurones artificiels	
II-1 : Généralité	19
II-1-1 : Fondement biologique	19
II-1-2 : Historique	20
II-2 : Neurone formel	20
II-1-1 : Représentation mathématique	21
II-1-2 : Fonctions d'activation	22
II-3 : Réseau de neurones artificiels	26
II-3-1 : Définition	26
II-3-2 : Architecture d'un réseau de neurones artificiels	26
II-3-2-1 : Réseau monocouche	26
II-3-2-2 : réseau multicouche	27
a) Réseaux multicouche	27
b) Réseaux multicouche classique	28
c) Réseaux a connexion locale	29
II-3-3 : L'apprentissage des réseaux.	29
II-3-3-1 : Apprentissage supervisé	30
II-3-3-2 : Apprentissage non supervisé	32
II-3-3-3 : Apprentissage hybride	33
II-3-3-4 : Règle d'apprentissage.	33

II-4 : Les perceptrons	34
II-4-1 : Définition	35
II-4-2 : Architecture	36
II-4-3 : algorithme d'apprentissage	36
II-5 : les cartes auto-organisatrices de Kohonen	38
II-4-1 : Définition	38
II-4-2 : Architecture	38
II-4-3 : algorithme d'apprentissage	39
II-6 : la mise en œuvre des réseaux de neurones artificiels	41
II-7 : La classification	42
II-8 : Conclusion	42
CHAPITRE III : Généralités sur les alliages d'aluminium	
III-1 : Introduction	43
III-2 : Les liaisons atomiques	43
III-3 : Structures cristallines	45
III-3-1 : Structure cubique centrée	45
III-3-2 : Structure cubique à faces centrées	46
III-3-3 : Structure hexagonale compacte	46
III-4 : Les alliages métalliques	47
III-4-1 : Formation de solutions solides	47
III-4-1-1 : Solutions solides d'insertion	47
III-4-1-2 : Solutions solides de substitution	48
III-4-1-3 : Règle de Hume-Rothery	49
III-5 : Les alliages d'aluminium	49
III-5-1 : L'aluminium	50
III-5-1-1 : Caractéristiques physiques	51
a) La masse volumique	51
b) Résistance à la corrosion	52
c) Conductivité électrique	52
III-5-1-2 : Propriétés mécaniques	52
III-5-2 : Les différents types d'alliage d'aluminium	53
III-5-3 : Alliages sans durcissement structural	55.
III-5-3-1 : La série 1000 (alliage non allié)	55
III-5-3-2 : La série 3000 (alliage Al-Mn)	55
III-5-3-3 : La série 3000 (alliage Al-Mn)	55
III-5-3-4 : Série 5000 (Al-Mg)	55
III-5-4 : Alliages à durcissement structural	56
III-5-4-1 : La série 2000 (alliage Al-Cu)	56
III-5-4-2 : La série 4000 (alliage Al-Si)	56
III-5-4-3 : Série 6000 (Al-Mg/Si)	57
III-5-4-4 : Série 7000 (Al-Zn-Mg et Al-Zn-Mg/Cu)	57
III-6 : Le durcissement structural	58
III-6-1 : Durcissement structural dans les alliages d'aluminium	59
III-6-2 : Exemple d'alliages de la série 7000	61

III-6-2-1 : Série 7005	61
III-6-2-2 : Série 7039	62
III-6-2-3 : Série 7075	62
III-6-2-4 : Série 7178	62
III-6-3 : Etat métallurgique des alliages d'aluminium	63
III-7 Conclusion	63
<i>Chapitre IV : Application des RNN à l'alliage Al-Zn</i>	
IV-1 : Introduction	64
IV-2 : La base de données	64
IV-3 : Le réseau de neurones	66
IV-4 : Architecture du réseau	68
IV-5 : Application à l'étude de l'alliage Al-Zn	69
IV-5-1 : Apprentissage du réseau	69
IV-5-2 : Présentation de la deuxième base d'exemple au réseau	72
IV-5-3 : Test de classification	75
IV-5-4 : Prédiction des propriétés élastique de l'alliage Al-Zn	79
IV-5-5 : Généralisation	83
IV-6 : Conclusion	85
IV-7 : Programme d'implémentation	86
Conclusion générale	89

Introduction générale

Alors que la quantité de données augmente de manière exponentielle, les gens sont en fait assoiffés de connaissance. La connaissance est obtenue par la compréhension de ces données.

Le data mining, un domaine de recherche particulièrement important de nos jours puisqu'il sert à fournir un sens aux données, passionnant, puisqu'il se base sur la capacité des ordinateurs à apprendre et à s'améliorer avec le temps, pour, par exemple, effectuer des prédictions.

Le *data mining*, textuellement minage de données mais souvent traduit en français par *fouille de données*, est un domaine qui consiste à comprendre les données, généralement par le moyen de méthodes statistiques. En d'autres termes, le *data mining* cherche à identifier des *tendances* parmi les données. Comme ce processus peut être très difficile, il est souvent comparé au minage de l'or dans les rivières: le gravier des alluvions représente l'énorme quantité de données et les pépites d'or représentent la connaissance cachées que l'on veut trouver.

En simplifiant légèrement, on peut diviser les différentes techniques de *data mining* suivant deux objectifs principaux: la prédiction et la description. La plupart des applications de *data mining* ont pour but la **prédiction**: pour un ensemble d'apprentissage dont on connaît les entrées et la sortie, on aimerait prédire la sortie d'un ensemble de test. Pour réaliser ces prédictions et extraire des informations, des connaissances que les données cachent, le *data mining* développe des méthodes et des outils comme la classification.

La classification est une étape importante pour une analyse de données efficace. Elle consiste à regrouper les objets d'un ensemble de données en classes homogènes. Il existe deux approches pour cette méthode, la classification supervisée et la classification non supervisée.

La classification supervisée est basée sur un ensemble de données appelé ensemble d'apprentissage dont la classe est connue. Le but étant de découvrir la structure des classes à partir de l'ensemble d'apprentissage afin de pouvoir généraliser cette structure sur un ensemble de données plus large.

L'outil ou l'apprentissage automatique utilisé pour la classification supervisée sont les réseaux de neurones artificiels et comme modèle le perceptron multicouche.

L'apparition de ces nouvelles méthodes numériques telles que les réseaux de neurones, la prévision et l'optimisation du comportement mécanique des matériaux sont devenues très accessibles. Ces méthodes sont devenues en quelques années des outils précieux dans les différents domaines de la science des matériaux, du fait de leur intérêt pratique qui est dû essentiellement à leurs propriétés mathématiques spécifiques et leurs mises en œuvre très simples.

Les scientifiques spécialisés dans l'étude des matériaux ont commencé à travailler avec ces méthodes d'exploration de données pour la conception de nouveaux matériaux, en se basant sur la compréhension des corrélations entre les différentes propriétés de ces matériaux.

Pour la conception des matériaux, ce n'est pas seulement l'obtention des données par des moyens théoriques, mais la manière de les analyser de façon efficace et compréhensible, et l'emploi de technique performante sont nécessaires. Ainsi dans l'étude des alliages métalliques la difficulté est l'analyse de grandes bases de données et de comprendre les corrélations entre différentes propriétés.

Les réseaux de neurones ont d'abord été développés pour résoudre des problèmes de contrôle, de reconnaissance de formes ou de mots, de mémorisation comme une alternative à l'intelligence artificielle, et en relation plus ou moins étroite avec la modélisation de processus cognitifs (capable de connaître ou faire connaître). Dans ce travail on va essayer d'explorer la puissance de cette technique du data mining, qui est les réseaux de neurones pour la classification de ces alliages suivant leurs propriétés mécaniques en vue de pouvoir prédire les propriétés d'un nouveau alliage.

Parmi les alliages métalliques, l'aluminium et ses alliages, jouent un rôle important dans l'industrie, ils occupent la deuxième place dans la construction de différentes pièces après le fer. Ces alliages, notamment les alliages à durcissement structural, sont de plus en plus utilisés dans l'industrie automobile et aéronautique du

fait de leur bon compromis densité/propriétés mécaniques (une faible densité d'aluminium et très bonnes propriétés mécaniques).

Actuellement les alliages ayant les meilleures caractéristiques mécaniques appartiennent à la série 7000, les principaux éléments d'addition sont dans ce cas ; le zinc dont la teneur varie de 4 à 8% massique, le magnésium dont la teneur varie de 1 à 3% massique et le cuivre. L'addition de cuivre, jusqu'à des teneurs de 2% massique, durcit ces alliages par solution solide, par ailleurs il diminue toute fois la trempabilité, la soudabilité et la ténacité de ces alliages, le magnésium et le chrome (<0,3%) améliorent leur résistance a la corrosion et la corrosion sous tension respectivement.

Le principal inconvénient de ces alliages est leur faible résistance à chaud, dès que la température de service dépasse 120 °C ; c'est pour cela que ces alliages sont très largement utilisés dans la construction aéronautique, et notamment pour la construction des avions civils et les tôles constituant les ailes d'avions.

Il est donc intéressant d'essayer d'étudier la ductilité de ces alliages en tenant compte du module de cisaillement et du module de compression ainsi que des différentes concentrations des éléments d'addition.

Le module de cisaillement (G) est pour un grand nombre de matériaux un indicateur de la dureté mécanique H. Le module de compression (B) représente la mesure de la rigidité moyenne de la liaison dans le cristal, il est proportionnel à l'énergie de cohésion. Le rapport B/G est un paramètre représentant la performance du matériau en terme de ductilité contre fragilité .La ductilité est caractérisé par un rapport B/G élevé (> 1.75) dans le cas contraire la valeur de B/G représente la fragilité du matériau.

La présente étude est divisée en quatre chapitres :

- Dans le chapitre I, nous rappelons les différentes définitions du data mining, avec un accent particulier sur les techniques et les méthodes employées pour la classification, qui connaissent actuellement un essor considérable dans l'analyse de données.

- Le chapitre II, est réservé à l'introduction aux réseaux de neurones ainsi qu'au principe de fonctionnement de cette technique .Nous résumons quelques modèles d'architecture de ces réseaux et leurs algorithmes d'apprentissage.
- Dans le chapitre III nous présentons des généralités sur les alliages d'aluminium et les différentes séries de ces alliages, en particulier l'alliage aluminium-zinc ainsi que leurs propriétés.
- Le chapitre IV est consacré à la mise au point d'un modèle de Réseaux de Neurones pour la classification de l'alliage (Al-Zn), en comparant les sorties désirées avec les sorties de ce réseau pour enfin la généralisation de ce modèle.
- Enfin, nous terminons cette étude par des conclusions et des perspectives.

Chapitre I

Datamining

I-1 Introduction

Le data mining, dans sa forme et sa compréhension actuelle, comme champ à la fois scientifique et industriel, est apparu au début des années 1990. Cette émergence n'est pas le fruit du hasard mais le résultat de la combinaison de nombreux facteurs à la fois technologiques, économiques et même sociopolitiques. On peut voir le « *data mining* » comme une nécessité imposée par le besoin des entreprises de valoriser les données qu'elles accumulent dans leurs bases.

En effet, le développement des capacités de stockage et les vitesses de transmission des réseaux ont conduit les utilisateurs à accumuler de plus en plus de données. Certains experts estiment que le volume des données double tous les ans. Que doit-on faire avec ces données coûteuses à collecter et à conserver ?

Le terme « *data mining* » est formé par analogie avec la mine, les mineurs passant au tamis d'énormes quantités de matériau sans valeur pour y découvrir de l'or. Appliqué au data mining, cet or représente les informations, jusque là inconnues ou indiscernables. Ces informations peuvent servir, par exemple, à déterminer des profils de client et d'achat type, les caractéristiques d'un segment de marché. [1]

I-2 Qu'est que le data mining

Le data mining est un procédé d'exploration et d'analyse de grands volumes de données en de les rendre compréhensibles. Il permet de découvrir des corrélations significatives en vu de trouver des règles de classement et de prédiction dont la finalité ultime est l'aide a la décision.

Grâce à des méthodes statistiques, mathématiques et de reconnaissance de formes le data mining est l'art d'extraire des connaissances a partir de données qui peuvent être stockées dans des entrepôts appelés « *data warehouse* ».

Le data mining permet de passer des données a la connaissance, il s'agit de fouilles visant à découvrir de l'information cachée que les données renferment.

C'est un mélange d'idées et d'outils provenant de la statistique, de l'intelligence artificielle et de l'informatique.

L'extraction automatique d'informations utiles, souvent jusque-là inconnue de grandes bases de données ou des ensembles de données afin d'aider à réduire, le modèle, de comprendre ou d'analyser les données. Les tâches prises en charge par l'exploration de données comprennent la prédiction, la segmentation, la modélisation de dépendance, le résumé, et le changement et la détection de déviation.

Le *data mining* est donc un domaine récent qui se situe à l'intersection des statistiques, de l'apprentissage automatique et des bases de données, figure I-1

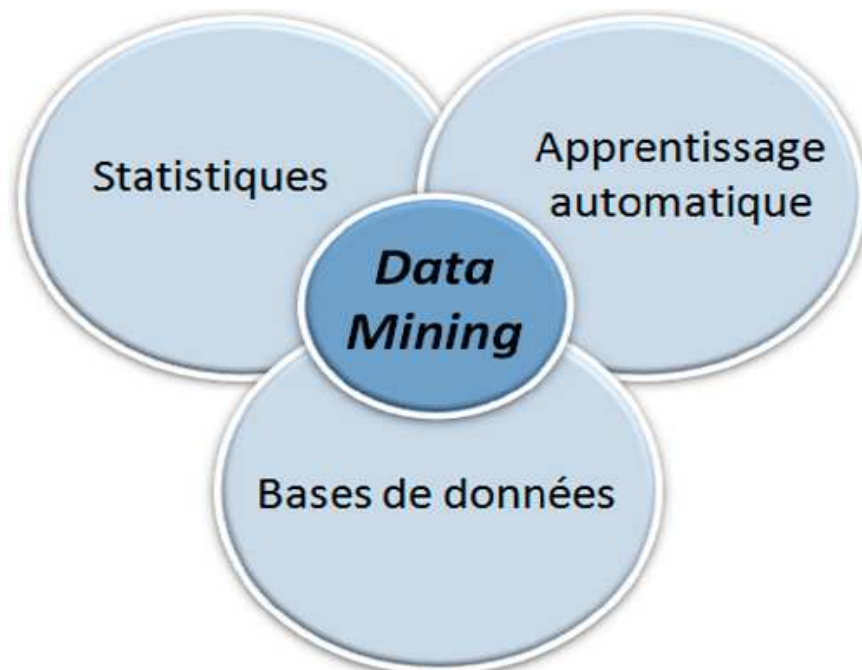


Figure I-1 Représentation schématique du data mining

I-3 Etapes d'un projet de data mining

Le data mining est un processus méthodique c'est une suite d'opérations ordonnées décrites comme suit :

Définition du problème : Quel est le but de l'analyse, que recherche-t-on ? Quels sont les objectifs ? Comment traduire le problème en une question pouvant servir de sujet d'enquête pour le data mining ?

A ce sujet, se souvenir que l'on travaille à partir des données existantes la question doit être ciblée selon les données disponibles. Cette phase consiste donc à comprendre et analyser les objectifs de l'application.

Collecte des données : Une phase absolument essentielle. On n'analyse que des données utilisables, c'est à dire "propres" et consolidées. On n'hésitera pas à extraire de l'analyse les données de qualité douteuse. Bien souvent, les données méritent d'être retravaillées. S'assurer au final que la quantité de données soit suffisante pour éviter de fausser les résultats. Créer une base de données pour la mise au point de l'application, cette phase de collecte nécessite le plus grand soin.

Construire le modèle d'analyse : ne pas hésiter à valider vos choix d'analyse sur plusieurs jeux d'essais en variant les échantillons. Une première évaluation peut nous conduire à reprendre les étapes (1 ou 2). Analyse statistique des données (réduction de la dimension), identifier le type de problèmes (discrimination, clustering) et choisir un algorithme.

Etude des résultats : Il est temps d'exploiter les résultats. Pour affiner l'analyse on n'hésitera pas à reprendre les étapes (1, 2 ou 3) si les résultats s'avéraient insatisfaisants, évaluer les performances de l'algorithme.

Formalisation et diffusion : Les résultats sont formalisés pour être diffusés. Ils ne seront utiles qu'une fois devenus une connaissance partagée. C'est bien là l'aboutissement de la démarche. C'est aussi là que réside la difficulté d'interprétation et de généralisation. [2]

I-4 Les tâches du data mining

On peut se poser la question suivante, que fait le data mining à la fin ?

On cite les principales tâches qui sont :

I-4-1 La classification

Elle consiste à examiner des caractéristiques d'un élément nouvellement présenté afin de l'affecter à une classe d'un ensemble prédéfini. Dans le cadre informatique, les éléments sont représentés par un enregistrement et le résultat de la classification viendra alimenter un champ supplémentaire.

La classification permet de créer des classes d'individus (terme à prendre dans son acception statistique). Celles-ci sont discrètes : homme / femme, oui / non, rouge / vert / bleu, ...

Les techniques les plus appropriées à la classification sont, les arbres de décision, le raisonnement basé sur la mémoire, éventuellement l'analyse des liens.

I-4-2 Clustering

L'analyse des clusters consiste à segmenter une population hétérogène en sous populations homogènes. Contrairement à la classification, les sous populations ne sont pas établis.

I-4-3 Règles d'associations

Une règle d'association est énoncée du type 60% des clients qui achètent du pain achètent aussi du lait .ces règles permettent de réorganiser les produits suivant le besoin des clients.[3]

I-4-4 La prédiction

La prédiction ressemble à la classification et à l'estimation mais dans une échelle temporelle différente. Tout comme les tâches précédentes, elle s'appuie sur le passé et le présent mais son résultat se situe dans un futur généralement précisé. La seule méthode pour mesurer la qualité de la prédiction est d'attendre !

Les techniques les plus appropriées à la prédiction sont Les arbres de décision les réseaux de neurones.

I-5 Les techniques du datamining

Au final que fait le data mining ?

Le data mining se base sur des techniques exploratrices utilisant les méthodes de l'apprentissage automatique représentation par clustering dans le cas de l'apprentissage non supervisée. Prédicatives utilisant l'apprentissage supervisé et l'explication par les règles d'association. C'est l'ensemble de règles du domaine des statistiques, des apprentissages automatiques et de l'optimisation en vu de prédire ou de décrire les phénomènes en instance. Ces techniques se divisent en deux méthodes.

I-5-1 Les techniques descriptives

Ces techniques visent à mettre en évidence des informations présentes mais cachées par le volume des données et il n'y a pas de variable cible « *prédire* ».

Parmi les méthodes, la classification hiérarchique a apprentissage non supervisée qui consiste à trouver dans l'espace de travail des groupes homogènes d'individus ou de variables utilisant les clusterings (segmentation) partitionnement logique de la base de données en clusters.

Dans ce cas de techniques comme seul les observations $\{x_1, x_2, \dots, x_n\}$ appartenant à un ensemble X^d sont disponibles, l'objectif est de décrire comment les données sont organisés et d'en extraire des sous-ensembles homogènes.

Exemple : On cherche à étudier le panier de la ménagère dans une certaine zone démographique en fonction de certains critères sociaux, (x) représente un individu à travers ces caractéristiques sociales et ses habitudes lors des courses.

Les techniques descriptives sont basées sur des modèles géométriques utilisant des algorithmes comme l'analyse en composante principale (ACP), l'analyse factorielle des correspondances, les méthodes de partitionnement (les centres noués, les k-means), classification neuronal (les cartes de Kohonen).

I-5-1-1 L'analyse en composantes principales ACP

C'est une technique descriptive pour l'exploitation des données, permettant d'étudier les relations qui existent entre les variables continues. Le but de l'ACP est d'identifier la structure de dépendance entre des observations multi variables afin d'obtenir une description ou une représentation compacte de ces dernières (variables) son utilisation a été restreinte à une projection des données sur différents axes factoriels.

Mathématiquement l'ACP est une technique de projection orthogonale linéaire qui projette les observations multidimensionnelles représentées dans un sous-espace de dimension m (m étant le nombre de variables observées) dans un sous-espace de dimension inférieur ($l < m$) en maximisant la variable. Elle peut être considérer comme une technique de minimisation de l'erreur quadratique d'estimation ou une technique de maximisation de la variance des projections (ces deux critères sont équivalents).[4]

I-5-1-2 Les k-means

C'est une des techniques de classification non supervisée (clustering) les plus utilisées. Etant donné un entier K , *K-means* partitionne les données en K groupes, ou "clusters", ou "classes" ne se chevauchant pas. Dans la méthode du *k-means*, une classe est représentée par son centre, qui est une moyenne (souvent pondérée) des objets appartenant à cette classe. [5]

Le résultat est obtenu en positionnant K "prototypes", ou "centroïdes" dans

les régions de l'espace les plus peuplées. Chaque observation est alors affectée au prototype le plus proche (règle dite "de la distance minimale"). Chaque classe contient donc les observations qui sont plus proches d'un certain prototype que de tout autre prototype.

Le principe du clustering est de diviser les données en plusieurs sous-ensembles. Comme le calcul de tous les sous-ensembles possibles n'est pas faisable, on utilise des heuristiques sous la forme d'optimisation itérative de la fonction objective.

I-5-1-3 La classification neuronale

Dans cette technique de classification non supervisée, sont utilisées les cartes auto-organisatrices de Kohonen ou cartes SOM (Self Organisation Maps). C'est des réseaux de neurones basés sur l'apprentissage non supervisée, qui établit une carte discrète, ordonnée topologiquement en fonction des entrées. Le réseau forme ainsi une sortie de treillis, où chaque nœud du treillis est un neurone associé à un vecteur de poids. La correspondance entre chaque vecteur de poids est calculée pour chaque entrée. Par la suite, le vecteur de poids ayant la meilleure corrélation, ainsi que certains de ses voisins vont être modifiés afin d'augmenter encore cette corrélation. Ceci est montré sur la figure I-2.

La carte se compose d'une grille régulière d'unités de traitement, "neurones". Un modèle d'une partie d'observation multidimensionnelle, éventuellement un vecteur consistant en des fonctionnalités, est associé à chaque unité. La carte vise à représenter toutes les observations disponibles avec une précision optimale en utilisant un ensemble restreint de modèles. Dans le même temps les modèles deviennent commandé sur la grille de sorte que des modèles semblables sont proches les uns des autres modèles et dissemblables loin les uns des autres.

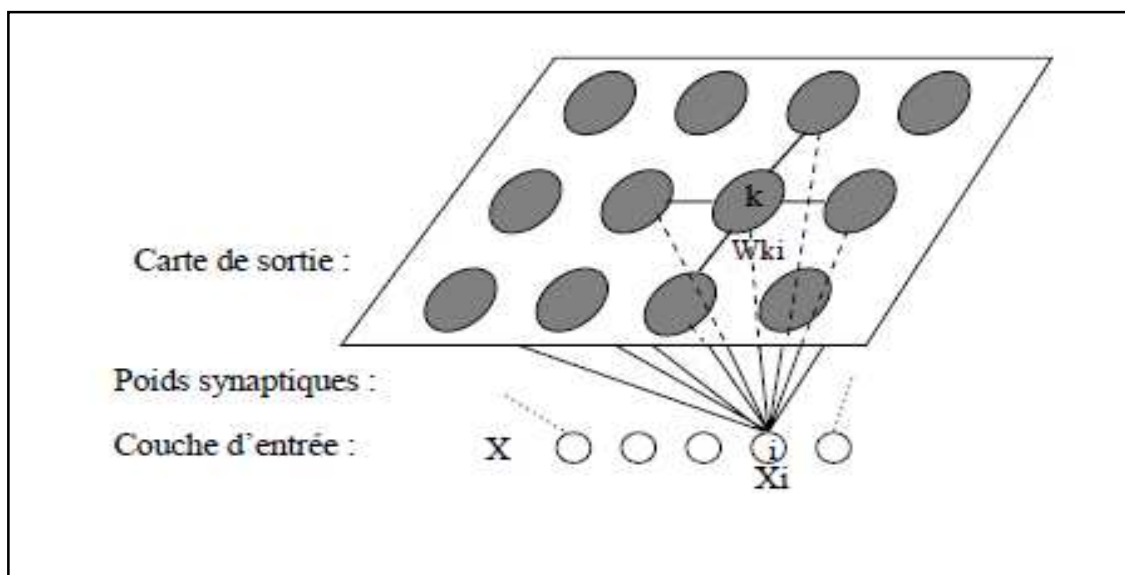


Figure I-2: Carte auto-organisatrice de Kohonen

Le SOM est un algorithme utilisé pour visualiser et interpréter de vastes ensembles de données de grande dimension. Les applications typiques sont la visualisation des états de processus ou les résultats financiers en représentant les dépendances au sein des centrales de données sur la carte.

Les cartes de Kohonen, appelées aussi cartes à préservation de proximité, disposent de capacités d'apprentissage automatique, d'un bon pouvoir illustratif des résultats et d'une forte sensibilité aux données fréquentes. Elles permettent également de réduire efficacement la complexité calculatoire. Le principe général de cet algorithme consiste à initialiser les poids des neurones de la couche de Kohonen en leur attribuant des valeurs aléatoires faibles, puis à chaque itération, présenter un vecteur de données au réseau, déterminer le neurone le plus proche à ce vecteur selon une mesure de distance et modifier les poids des neurones appartenant au voisinage du neurone élu.

Ainsi, l'algorithme d'apprentissage de la carte topologique comprend principalement deux étapes : la première consiste à sélectionner un neurone gagnant et la seconde consiste à mettre à jour le poids de ce neurone et des neurones de son voisinage.

I-5-1-4 Domaines d'application

- *Marketing* : segmentation du marché en découvrant des groupes de clients distincts à partir de bases de données d'achats.
- *Environnement* : identification des zones terrestres similaires(en termes d'utilisation) dans une base de données d'observation de la terre.
- *Assurance* : identification de groupes d'assurés distincts associés à un nombre important de déclarations.
- *Planification de villes* : identification de groupes d'habitations suivant le type d'habitation, valeur, localisation géographique, ...
- *Médecine*: Localisation de tumeurs dans le cerveau (nuage de points du cerveau fournis par le neurologue).
- La reconnaissance automatique de la parole
- Analyse de la voix clinique
- Suivi de l'état des plantes et des procédés industriels
- Classification des nuages à partir d'images satellites
- Analyse des signaux électriques en provenance du cerveau
- Organisation de la recherche et de grandes collections de documents (les **WEBSOM** méthode)
- Analyse et visualisation des grandes collections de données statistiques (par exemple, les données macro-économiques)

I-5-2 Les techniques prédictives

C'est des méthodes servant à expliquer et prévoir un ou plusieurs phénomènes observables et effectivement mesurés. L'analyse se fait sur l'ensemble des variables de la base de données utilisant une classification à apprentissage supervisée (la classe est connue) basée sur l'apprentissage automatique. Ces techniques visent à extrapoler de nouvelles informations à partir des informations cachées (c'est le cas des scoring), dans ce cas il y a une variable cible à prédire.

Les méthodes prédictives utilisent des modèles à base de règles logiques (les arbres de décision), des modèles à base de fonctions mathématiques comme les réseaux de neurones à apprentissage supervisé (le perceptron multicouches) ou la régression linéaire (régression PLS), les algorithmes génétiques, les règles d'associations.

L'objectif de l'apprentissage supervisé est de partir d'un ensemble d'observation $\{x_1, x_2, \dots, x_n\}$ appartenant à un ensemble X^d et des mesures $\{y_i\}$ appartenant à Y , on cherche à estimer les dépendances entre l'ensemble X et Y .

Exemple : on cherche à estimer les liens entre les habitudes alimentaires et le risque d'infarctus, (x_i) est un patient décrit par (d) caractéristiques concernant son régime et (y_i) une catégorie (risque, pas risque). On parle d'apprentissage *supervisé* car les (y_i) permettent de guider le processus d'estimation.

I-5-2-1 Les arbres de décision

La popularité de la méthode repose en grande partie sur sa simplicité. Il s'agit de trouver un partitionnement des individus que l'on représente sous la forme d'un arbre de décision. Construction du plus petit arbre de décision possible, une branche pour chaque valeur d'un attribut, les feuilles désignent la classe de l'objet à classer

C'est une structure de données utilisée comme modèle pour la classification des sous-groupes. L'objectif est de produire des groupes d'individus les plus homogènes possibles du point de vue de la variable à prédire. Il est d'usage de représenter la distribution empirique de l'attribut à prédire sur chaque sommet (nœud) de l'arbre.

Exemple : [6]

Etant donné un ensemble de 14 observations, il s'agit d'expliquer le comportement des individus par rapport à un jeu {jouer, ne pas jouer} à partir des prévisions météorologiques (soleil, couvert, pluie).

Par exemple, l'arbre de la figure I-3, ci-dessus décide une réponse booléenne classification dans l'ensemble (oui, non) en fonction des valeurs discrètes des attributs (difficile, durée, motivation, surprenant).

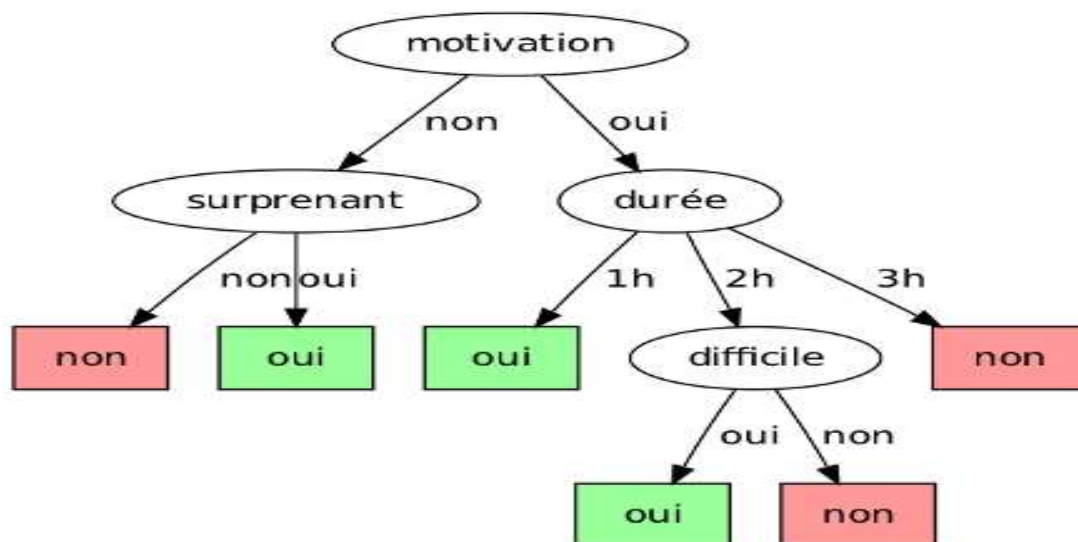


Figure I-3:Exemple d'arbre de décision

I-5-2-2 La classification supervisée

Etant donné un ensemble d'individus dont la classe est connue, le but étant de découvrir la structure des classes à partir de l'ensemble afin de pouvoir généraliser cette structure sur un ensemble de données plus large. Dans le but de la prédiction, les réseaux de neurones ont été appliqués avec succès à l'apprentissage de tâches de classification et d'approximation de fonctions. C'est un est le processus d'adaptation des paramètres d'un système pour donner une réponse désirée à une entrée ou stimulation quelconque.

Le perceptron multicouche est un type de réseau de neurones artificiels dont l'objectif de son utilisation comme outil de data mining est son algorithme d'apprentissage supervisé [7].

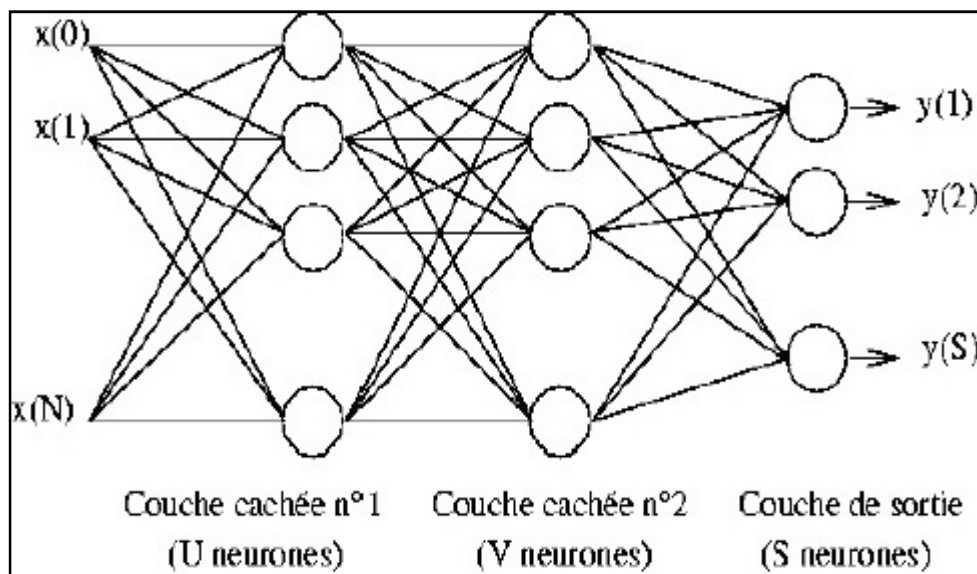


Figure I-4 Perceptron multicouches

Nombreux domaines sont les applications de ces méthodes à cause de leur algorithme d'apprentissage supervisé plus particulièrement le perceptron multicouche. On cite les grands domaines qui sont la classification, l'approximation des fonctions, regroupement, la prédiction, l'optimisation et la commande robotique. Ces méthodes sont appliquées dans les domaines de finance comme l'accord de crédit, marketing ciblé, diagnostic médical (analyse de l'effet d'un traitement), détection de fraudes fiscales.

Exemple de reconnaissance de chiffre sur la figure I-5

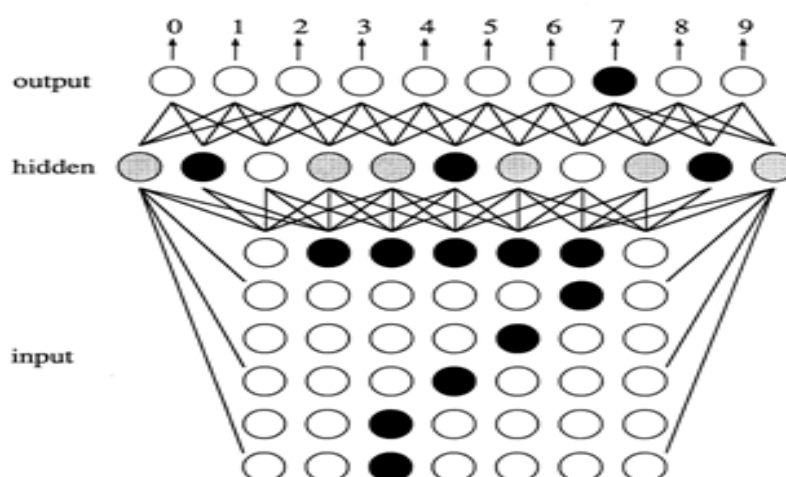


Figure I-5 Application des réseaux de neurones pour la reconnaissance des chiffres

I-6 Conclusion

Le data mining est une discipline née en dehors de la statistique, dans la communauté des bases de données et de l'intelligence artificielle, cherchant à découvrir des structures inconnues et utiles à l'aide de processus de recherche algorithmique, linéaire ou non, à la fin de tout le processus, le but du data mining est d'aider à prendre des décisions, après une analyse rationnelle. [8]

On s'aperçoit donc que les divers domaines d'applications du data mining les plus connues sont la reconnaissance d'images, de textes et du langage, utilisation dans l'analyse de séquences ADN, finance, le Web, la reconnaissance du *spam*. Il est clair que ce domaine en pleine croissance va intéresser la plupart des entreprises qui travaillent avec de grandes quantités de données. Quelle entreprise ne voudrait pas de prédictions sur la consommation de ses clients ou de recommandations automatiques pour que le client achète plus?

Le data mining est utile et utilisé dans un grand nombre de domaines, pour des problèmes variés, et les outils employés pour l'exploitation et l'exploration de grandes bases de données sont nombreux.

En partant de cette idée que le data mining est utilisé même dans la recherche scientifique et plus précisément dans la prédiction des nouveaux matériaux. En effet l'étude des matériaux se base sur l'étude de leur propriétés physiques, mais quelle technique et méthode adopter pour s'approcher des résultats expérimentaux, en employant l'outil informatique et prédire de nouveaux matériaux.

Dans le chapitre suivant on essaye de présenter une des techniques prédictives qui est les réseaux de neurones artificiels et plus précisément un algorithme d'apprentissage pour une classification supervisée « *Perceptron multicouche* ».

CHAPITRE II
Les Réseaux de neurones artificiels

II-1 Généralités

Les réseaux de neurones artificiels sont devenus en quelques années des outils précieux dans des domaines très divers de l'industrie et divers services, des techniques d'analyse en provenance de l'apprentissage automatique (intelligence artificielle) et des bases de données. Néanmoins, ils n'ont pas encore atteint leur plein développement pour des raisons plus psychologiques que techniques, liées aux connotations biologiques du terme et au fait qu'ils sont considérés, à tort, comme des outils d'Intelligence Artificielle. Or l'intérêt des réseaux de neurones, dans le domaine des Sciences ne doit rien à la métaphore biologique, il est uniquement dû aux propriétés mathématiques spécifiques de ces réseaux. Les réseaux de neurones ont d'abord été développés pour résoudre des problèmes de contrôle, de reconnaissance de formes ou de mots, de décision, de mémorisation comme une alternative à l'intelligence artificielle, et en relation plus ou moins étroite avec la modélisation de processus cognitifs (capable de connaître ou faire connaître) réels et des réseaux de neurones biologiques. [11]

II-1-1 Fondement biologique

Dans le cerveau humain, un neurone typique recueille les signaux des autres à travers une multitude de structures fines appelées dendrites. Le neurone envoie des pics d'activité électrique à travers un long, un support mince connu sous le nom d'un axone, qui se divise en des milliers de branches. A la fin de chaque branche, une structure appelée synapse convertit l'activité de l'axone en compte les effets électriques qui inhibent ou excitent l'activité de l'axone en compte les effets électriques qui inhibent ou excitent l'activité dans les neurones connectés. Quand un neurone reçoit une entrée excitatrice qui est suffisamment grande par rapport à son entrée inhibitrice, il envoie un pic d'activité électrique vers le bas de son axone.

II-1-2 Historique

De façon générale, on situe le début des réseaux de neurones artificiels en 1943, avec les travaux de « *McCulloch et Pitts* » qui montrent qu'un réseau de neurones discret, sans contrainte de topologie, peut représenter n'importe quelle fonction booléenne et donc émuler un ordinateur.

En 1958, « *Rosenblatt* » propose le premier algorithme d'apprentissage, qui permet d'ajuster les paramètres d'un neurone. En 1969, Minsky et Papert publient le livre 'Perceptrons', dans lequel ils utilisent une solide argumentation mathématique pour démontrer les limitations des réseaux de neurones à une seule couche.

En 1982, « *Hopfield* » propose des réseaux de neurones associatifs et l'intérêt pour les réseaux de neurones renaît chez les scientifiques.

En 1986, « *Rumelhart, Hinton et Williams* » publient l'algorithme de la 'rétropropagation de l'erreur' qui permet d'optimiser les paramètres d'un réseau de neurones à plusieurs couches. À partir de ce moment, la recherche sur les réseaux de neurones connaît un essor fulgurant et, au cours des années 90, les applications commerciales de ce succès académique suivent. [12]

II-2 Neurone formel

Le neurone artificiel (ou cellule) est un processeur élémentaire. Il reçoit un nombre variable d'entrées en provenance de neurones appartenant à un niveau situé en amont (on parlera de neurones "amont"). A chacune des entrées est associé un poids w représentatif de la force de la connexion.

Chaque processeur élémentaire est doté d'une sortie unique, qui se ramifie ensuite pour alimenter un nombre variable de neurones appartenant à un niveau situé en aval (on parlera de neurones "avals"). A chaque connexion est associé un poids (figure II-1).

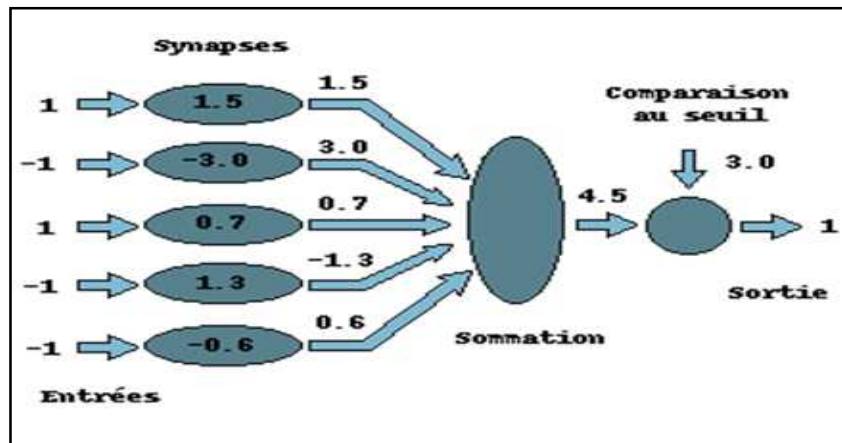


Figure II-1 Neurone de Mac Culloch et Pitts

II-2- 1 Représentation mathématique

Les réseaux de neurones formels sont à l'origine d'une tentative de modélisation mathématique du cerveau humain (figure II-2). Les premiers travaux datent de 1943 et sont l'œuvre de MM. Mac Culloch et Pitts. Ils présentent un modèle assez simple pour les neurones et explorent les possibilités de ce modèle. La modélisation consiste à mettre en œuvre un système de réseau neuronal sous un aspect non pas biologique mais artificiel. Un neurone est donc avant tout un opérateur mathématique, dont on peut calculer la valeur numérique par quelques lignes de logiciel. D'un point de vue mathématique, un neurone formel réalise donc une fonction linéaire. [13]

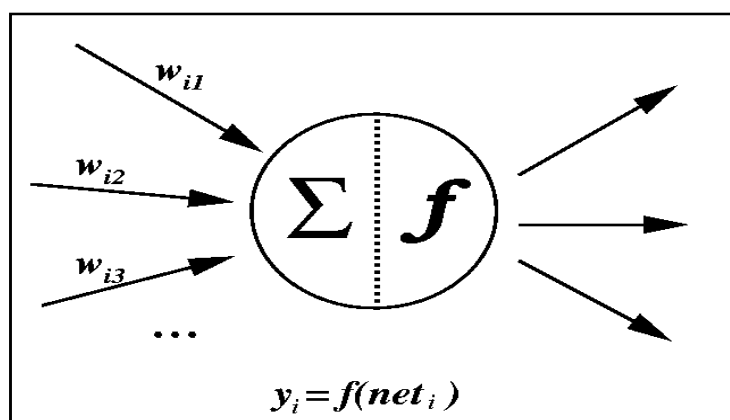


Figure II. 2 Neurone formel

$$y = f(x_1, x_2, \dots, x_p; w_1, w_2, \dots, w_p)$$

Les $\{x_i\}$ sont les variables (ou entrées), les $\{w_j\}$ sont des paramètres (ou poids).

Les entrées peuvent être booléennes, binaires (0,1), bipolaire (-1,1) ou réels.

En règle générale, le calcul de la valeur de cette fonction peut se décomposer en deux étapes :

- Une combinaison linéaire des entrées

$$U = \sum_{i=1}^n w_i x_i - \theta$$

- La sortie du neurone est :

$$y = f(U) = f(\sum_{i=1}^n w_i x_i - \theta)$$

(θ) est appelé biais, il peut être considéré comme la pondération de l'entrée 0 fixée à 1

(U) est appelé potentiel du neurone.

La fonction (f) permet de définir l'état interne du neurone en fonction de son entrée totale.

Citons à titre d'exemple quelques fonctions d'activation ou de transfert souvent utilisées .

II-2-2 Fonctions d'activation

Dans sa première version, le neurone formel était donc implémenté avec une fonction à seuil, mais de nombreuses versions existent. Ainsi le neurone de McCulloch et Pitts a été généralisé de différentes manières, en choisissant d'autres fonctions d'activations. [14]

La sortie du neurone dépend d'une fonction de transfert, dont les principales sont :

a) fonction binaire à seuil

$$h(x) = \begin{cases} 1 & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases} \quad \text{sgn}(x) = \begin{cases} -1 & \text{si } x < 0 \\ 1 & \text{sinon} \end{cases}$$

Le seuil introduit une non-linéarité dans le comportement du neurone, cependant il limite la gamme des réponses possibles à deux valeurs.

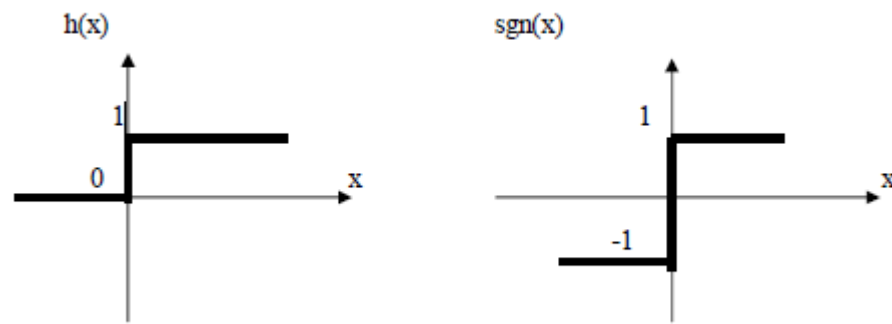


Figure II-3 Fonctions binaires a seuil

b) Fonction linéaire

C'est l'une des fonctions d'activations les plus simples, sa fonction est définie par : $F(x)=x$

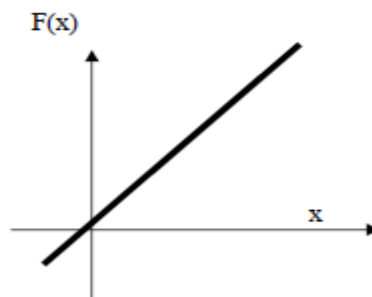


Figure II-4 Fonction linéaire

c) Fonction linéaire à seuil ou multi seuil

Cette fonction représente un compromis entre la fonction linéaire et la fonction seuil, entre ses deux barres de saturation, elle confère au neurone une gamme de réponses possibles. En modulant la pente de la linéarité, on affecte la plage de réponse du neurone.(figure II-5)

$$F(x)=\begin{cases} x & x \in [u, v] \\ v & \text{si } x \geq v \\ u & \text{si } x \leq u \end{cases}$$

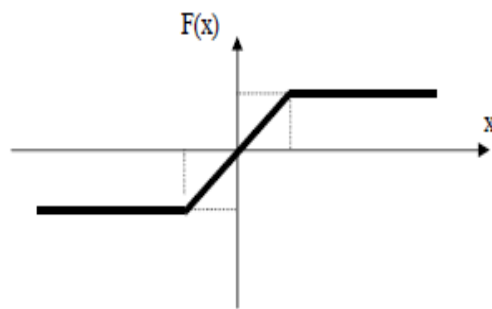


Figure II-5 Fonction linéaire a seuil

d) La fonction sigmoïde

Elle est l'équivalent continu de la fonction linéaire. Etant continu, elle est dérivable, d'autant plus que sa dérivée est simple à calculer, (figure II-6) elle est définie par :

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

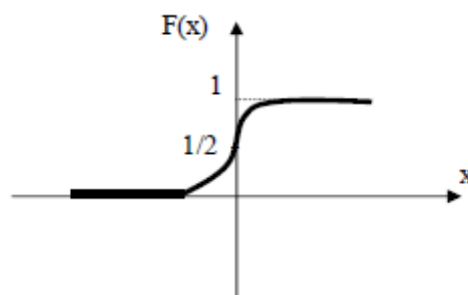


Figure II-6 Fonction sigmoïde

Un neurone formel ne réalise donc rien d'autre qu'une somme pondérée suivie d'une non linéarité. C'est l'association de tels éléments simples sous la forme de réseaux qui permet de réaliser des fonctions utiles pour des différentes applications.

Les fonctions réalisées par un neurone formel sont représentées sur les figures (II-7), (II-8) et (II-9)

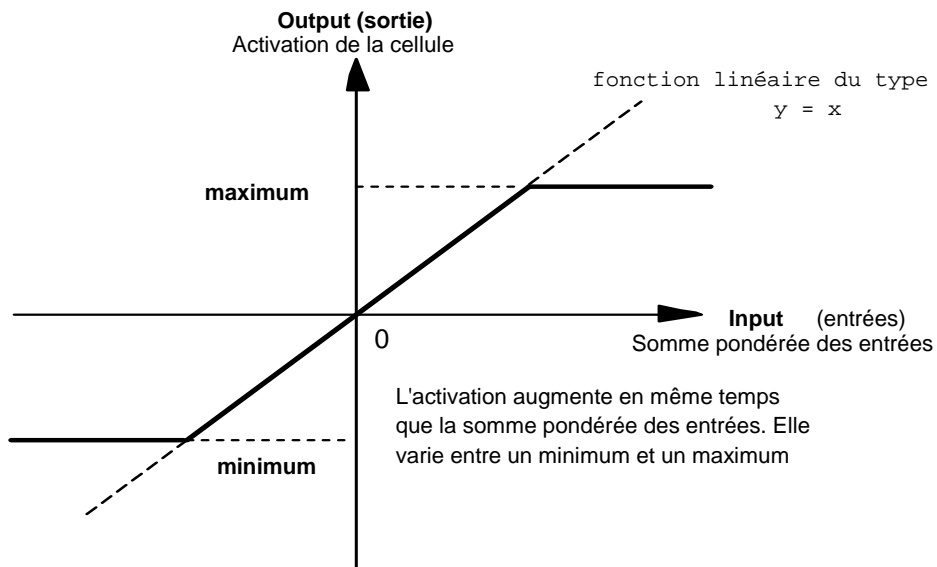


Figure II-7 : Fonction linéaire

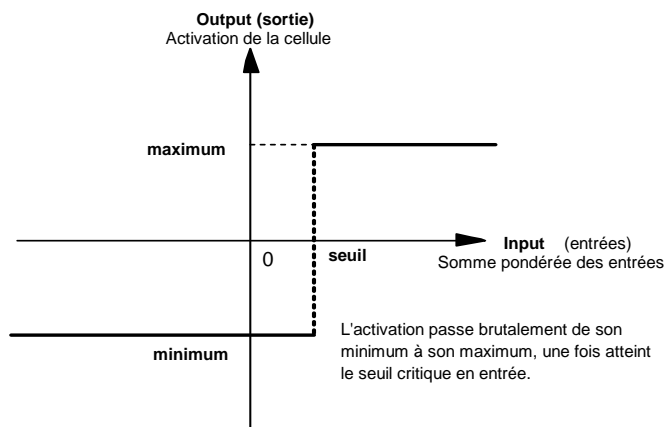


Figure II-8 : Fonction à seuil

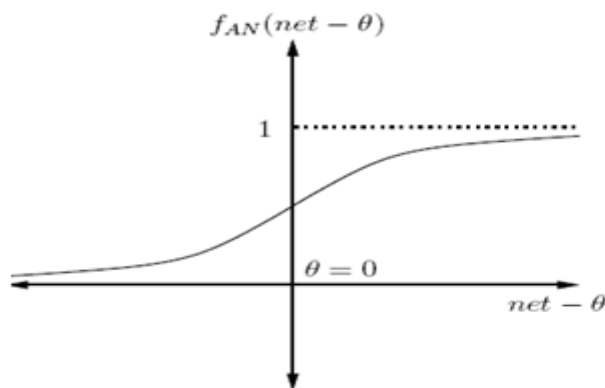


Figure II-9: fonction sigmoïde

II-3 Réseau de neurones artificiel (RNN)

II-3-1 Définition

Un ANN (Réseau de Neurones Artificiels) est un ensemble de neurones formels (d'unités de calcul simples, de nœuds processeurs) associés en couches (ou sous-groupes) et fonctionnant en parallèle.

Dans un réseau, chaque sous-groupe fait un traitement indépendant des autres et transmet le résultat de son analyse au sous-groupe suivant. L'information donnée au réseau va donc se propager couche par couche, de la couche d'entrée à la couche de sortie, en passant soit par aucune, soit par plusieurs couches intermédiaires (dites couches cachées).

Habituellement (excepté pour les couches d'entrée et de sortie), chaque neurone dans une couche est connecté à tous les neurones de la couche précédente et de la couche suivante. Les RNA ont la capacité de stocker de la connaissance empirique et de la rendre disponible à l'usage. Les habiletés de traitement (et donc la connaissance) du réseau vont être stockées dans les poids synaptiques, obtenus par des processus d'adaptation ou d'apprentissage. [15]

II-3-2 Architecture d'un réseau de neurones artificiel

Les connexions entre les neurones qui composent le réseau décrivent la topologie du modèle. Elle peut être quelconque, mais le plus souvent il est possible de distinguer une certaine régularité (réseau à connexion complète).

II-3-2-1 Réseaux monocouche

La structure d'un réseau monocouche est telle que des neurones organisés en entrée soient entièrement connectés à d'autres neurones organisés en sortie par une couche modifiable de poids.

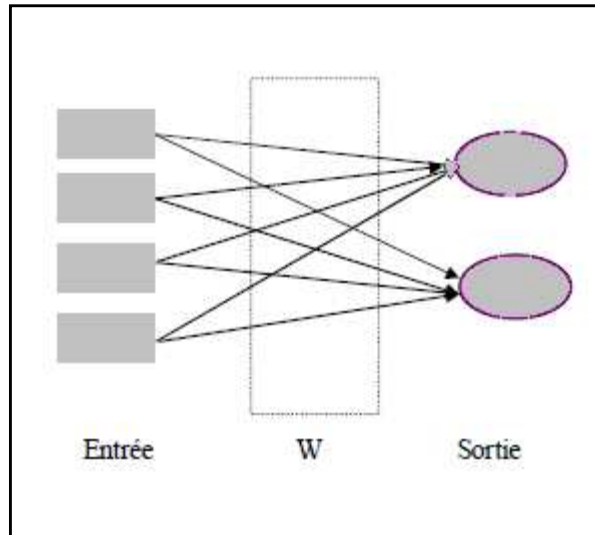


Figure II-10 réseau de neurones monocouche

II-3-2-2 Réseaux multicouche

Les neurones sont arrangés par couche. Il n'y a pas de connexion entre neurones d'une même couche, et les connexions ne se font qu'avec les neurones de couches avales. Habituellement, chaque neurone d'une couche est connecté à tous les neurones de la couche suivante et celle-ci seulement. Ceci nous permet d'introduire la notion de sens de parcours de l'information (de l'activation) au sein d'un réseau et donc définir les concepts de neurone d'entrée, neurone de sortie. Par extension, on appelle couche d'entrée l'ensemble des neurones d'entrée, couche de sortie l'ensemble des neurones de sortie. Les couches intermédiaires n'ayant aucun contact avec l'extérieur sont appelées couches cachées.

a) Réseau multicouche classique :Figure (II- 11)

Dans un réseau multicouche classique, il n'y a pas de connexion entre neurones d'une même couche et les connexions ne se font qu'avec les neurones de la couche aval. Tous les neurones de la couche amont sont connectés à tous les neurones de la couche aval.

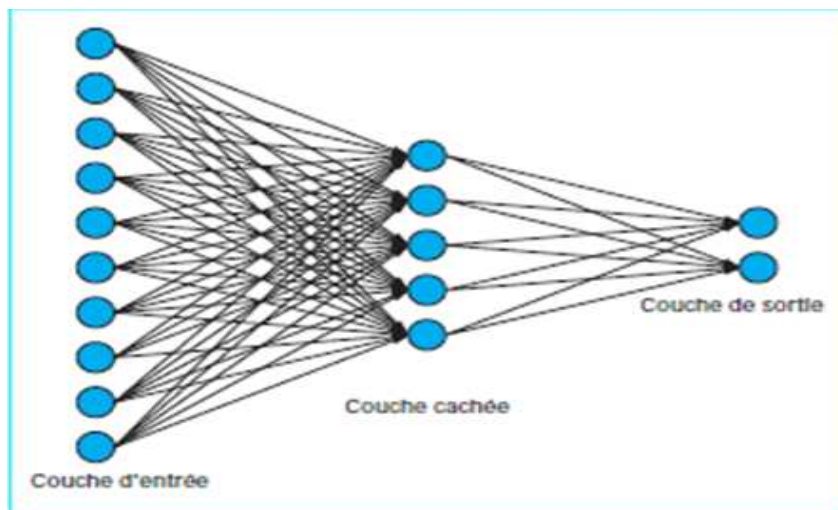


Figure II-11 Réseau multicouche classique

b) Réseau à connexions locales : figure (II-12)

C'est aussi un réseau multicouche, mais tous les neurones d'une couche amont ne sont pas connectés à tous les neurones de la couche aval. Nous avons donc dans ce type de réseau de neurones un nombre de connexions moins important que dans le cas du réseau de neurones multicouche classique.

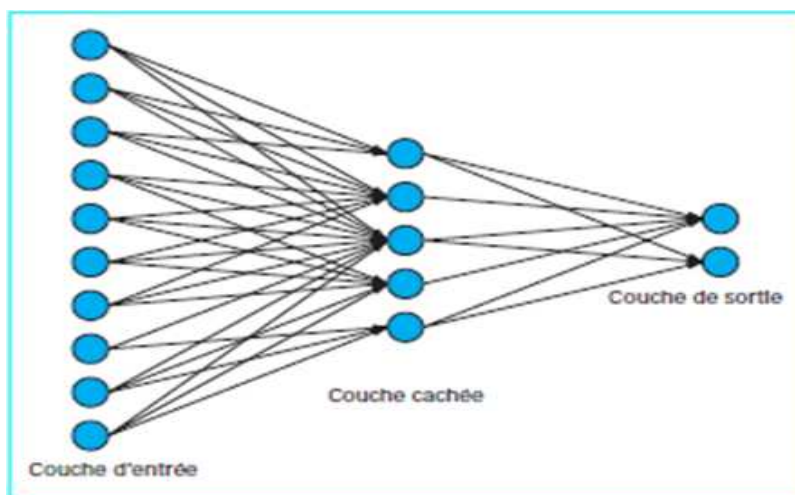


Figure II-12 Réseau à connexion locale

c) Réseau à connexions récurrentes : figure (II-13)

Un réseau de ce type signifie qu'une ou plusieurs sorties de neurones d'une couche aval sont connectées aux entrées des neurones de la couche amont ou de la même couche. Ces connexions récurrentes ramènent l'information en arrière par rapport au sens de propagation défini dans un réseau multicouche. Les réseaux à connexions récurrentes sont des réseaux plus puissants car ils sont séquentiels plutôt que combinatoires comme l'étaient ceux décrits précédemment. La rétroaction de la sortie vers l'entrée permet à un réseau de ce type de présenter un comportement temporel.

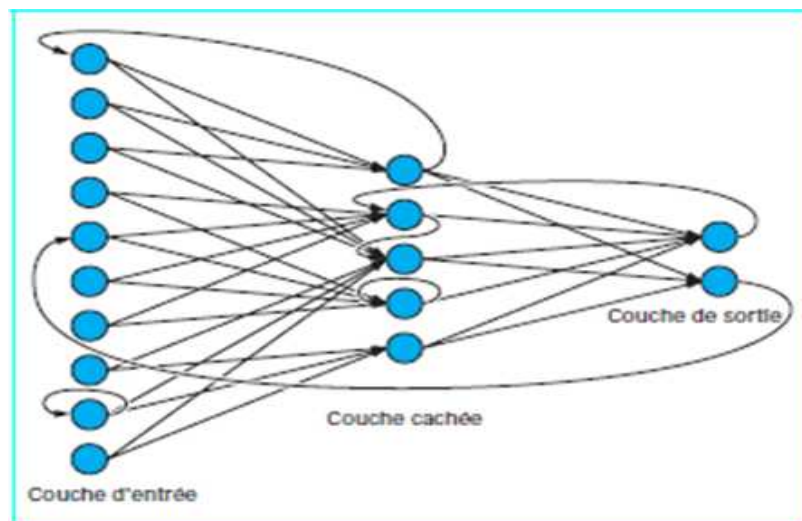


Figure II-13 Réseau à connexions récurrentes

II-3-3 L'apprentissage des réseaux de neurones

La recherche des algorithmes d'apprentissage statistiques a mené à d'importantes découvertes au cours des deux dernières décennies, qui ont changé la manière de concevoir le problème de rendre les ordinateurs plus intelligents. Un agent est intelligent parce qu'il a des connaissances opérationnelles (pas nécessairement sous une forme verbale explicite) qui lui permettent d'effectuer certaines tâches ou de répondre à certaines questions sur un certain domaine.

L'objectif de l'apprentissage est de fournir une méthode au réseau, afin qu'il puisse ajuster ses paramètres lorsqu'on lui présente des exemples à traiter. On distingue habituellement trois types d'apprentissage : supervisé, non supervisé et hybride.

En effet les réseaux de neurones ont été appliqués avec succès à l'apprentissage de tâches de classification et d'approximation de fonctions. L'apprentissage est le processus d'adaptation des paramètres d'un système pour donner une réponse désirée à une entrée ou stimulation quelconque.

II-3-3-1 Apprentissage supervisé : figure (II-14)

Dans l'apprentissage supervisé, on présente à l'ordinateur des exemples sous la forme de paires (entrée, sortie désirée).

Dans ce type d'apprentissage, on fournit au réseau de neurones la donnée à traiter mais aussi la sortie attendue. Le réseau effectue une évaluation de la donnée, puis compare la valeur obtenue avec la valeur désirée, il va ensuite modifier ses paramètres internes afin de minimiser l'erreur constatée. L'apprentissage 'supervisé' consiste donc à calculer les coefficients de telle manière que les sorties du réseau de neurones soient, pour les exemples utilisés lors de l'apprentissage, aussi proches que possibles des sorties 'désirées', qui peuvent être : la classe d'appartenance de la forme que l'on veut classer, la valeur de la fonction que l'on veut approcher ou de la sortie du processus que l'on veut modéliser, ou encore la sortie souhaitée du processus à commander. Il est habituel de le présenter en utilisant le paradigme du professeur et de l'élève.

De façon conceptuelle, on admet qu'il existe un professeur qui connaît la relation exacte entre toutes les entrées et leurs sorties, mais le réseau (élève) ne connaît pas cette relation. Si l'élève et le professeur sont exposés à une même entrée, le professeur est capable d'indiquer à l'élève la réponse désirée. Les paramètres (dans le cas des réseaux de neurones, le nombre de neurones et les poids) de l'élève doivent être ajustés pour donner la même réponse que celle du professeur. Cet ajustement, l'apprentissage, est réalisé en général de façon itérative, en minimisant une mesure de l'erreur, jusqu'à ce que le réseau-élève puisse émuler aussi bien que possible le professeur.

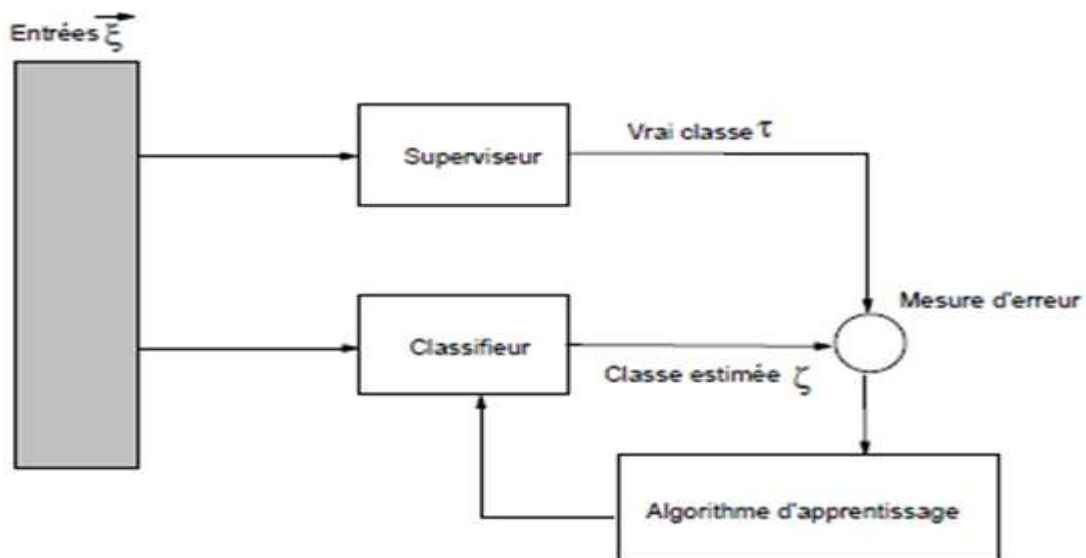


Figure II-14 réseau a apprentissage supervisé

La plupart des algorithmes d'apprentissage des réseaux de neurones sont des algorithmes d'optimisation : ils cherchent à minimiser, par des méthodes d'optimisation non linéaire, une fonction de coût, qui constitue une mesure de l'écart entre les réponses réelles du réseau et ses réponses désirées.

Cette optimisation se fait de manière itérative, en modifiant les poids en fonction du gradient de la fonction de coût : le gradient est estimé par une méthode spécifique aux réseaux de neurones, dite méthode de rétro propagation, puis il est utilisé par l'algorithme d'optimisation proprement dit. Les poids sont initialisés aléatoirement avant l'apprentissage, puis modifiés itérativement, jusqu'à obtention d'un compromis satisfaisant entre la précision de l'approximation sur l'ensemble d'apprentissage et la précision de l'approximation sur un ensemble de validation, distinct du précédent. Contrairement à des affirmations maintes fois répétées, l'apprentissage des réseaux de neurones n'est pas spécialement lent : il existe des algorithmes d'optimisation non linéaire extrêmement rapides qui permettent de faire des développements industriels sur de simples PC. Le but fondamental de l'apprentissage est de bien généraliser à de nouveaux cas. [16]

II-3-3-2 Apprentissage non supervisé

Dans ce type d'apprentissage, aucune information n'est fournie au réseau en plus des données à apprendre. Celui-ci est amené à découvrir la structure sous-jacente des données afin de les organiser en clusters. L'apprentissage non-supervisé correspond au cas où aucune cible n'est prédéterminée.

Ainsi, l'ensemble d'entraînement ne contient que des entrées et ne définit pas explicitement la nature de la fonction (f) qui doit être retournée par l'algorithme d'apprentissage. C'est plutôt l'utilisateur qui doit spécifier le problème à résoudre.

Pour ce problème, (f) doit fournir une estimation de la fonction de densité ou de probabilité de la distribution ayant généré les éléments d'entrées. Pour le problème de la classification les cartes auto-organisatrices de Kohonen sont utilisées dans les réseaux de neurones artificiels.

Dans l'apprentissage non supervisé, le réseau modifie ses paramètres en tenant compte seulement des informations locales. Ces méthodes n'ont pas besoins de sorties désirées préétablies. Les réseaux utilisant cette technique sont appelés réseaux à dynamique autonome et sont considérés comme des détecteurs de régularité, car le réseau apprend en détectant les régularités dans la structure des motifs d'entrée et produit la sortie la plus satisfaisante. [17]

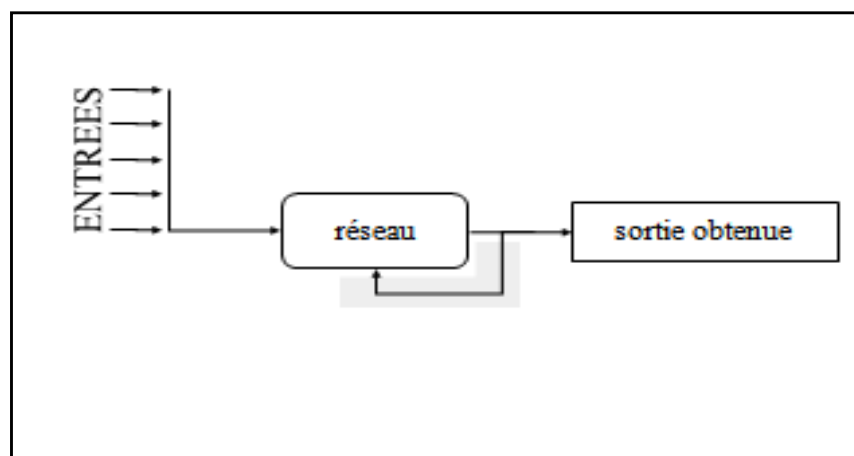


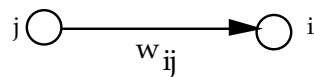
Figure II-15 réseau a apprentissage non supervisé

II-3-3-3 Apprentissage hybride

Ce type d'apprentissage est plus rare et est encore mal exploré, cette approche combine des **méthodes** numériques (réseaux de neurones, algorithmes génétiques) et des méthodes symboliques. Certains auteurs utilisent le terme d'apprentissage hybride pour parler d'un couplage 'supervisé non supervisé' ; dans ce cas, il s'agit d'un réseau qui met en parallèle ou en série un réseau entraîné en mode supervisé et un autre en mode non supervisé.

II-3-3-4 Règles d'apprentissage des réseaux de neurones artificiels

L'apprentissage est une phase du développement d'un réseau de neurones durant laquelle le comportement du réseau est modifié jusqu'à l'obtention du comportement désiré. Cette modification consiste à modifier les poids des connexions entre les neurones en suivant certaines. La loi de Hebb (1949) s'applique aux connexions entre neurones, comme le représenté :



(i) le neurone amont, (j) le neurone aval et (w_{ij}) le poids de la connexion.

Elle s'exprime de la façon suivante :

"Si 2 cellules sont activées en même temps alors la force de la connexion augmente".

La modification de poids dépend de la coactivation des neurones pré synaptique et post synaptique.

x_i et x_j sont respectivement les valeurs d'activation des neurones i et j .

∂w_{ij} (dérivée partielle du poids) correspond à la modification de poids réalisée.

La loi de Hebb peut être modélisée par les équations suivantes :

$w(t+1)$ est le nouveau poids

$w_{ij}(t)$ l'ancien

$$w_{ij}(t+1) = w_{ij}(t) + \partial w_{ij}(t)$$

$$\partial w_{ij}(t) = x_i \cdot x_j \text{ (l'activité est modélisée comme le produit des deux valeurs d'activation)}$$

Algorithme d'apprentissage

L'algorithme d'apprentissage modifie de façon itérative (petit à petit) les poids pour adapter la réponse obtenue à la réponse désirée. Il s'agit en fait de modifier les poids lorsqu'il y a erreur seulement.

1/ Initialisation des poids et du seuil S à des valeurs (petites) choisies au hasard.

2/ Présentation d'une entrée $E_i = (e_1, \dots, e_n)$ de la base d'apprentissage.

3/ Calcul de la sortie obtenue x pour cette entrée : $a = \sum (w_i \cdot e_i) - S$

la valeur de seuil est introduite ici dans le calcul de la somme pondérée

$$x = \text{signe}(a) : \begin{cases} \text{si } a > 0 \text{ alors } x = +1 \\ \text{sinon } a \leq 0 \text{ alors } x = -1 \end{cases}$$

4/ Si la sortie (x) est différente de la sortie désirée (d) pour cet exemple d'entrée E_i alors modification des poids (μ est une constante positive, qui spécifie le pas de modification des poids) : $w_{ij}(t+1) = w_{ij}(t) + \mu \cdot (x_i \cdot x_j)$

5/ Tant que tous les exemples de la base d'apprentissage ne sont pas traités correctement (i.e. modification des poids), retour à l'étape 2.

II-4 Les perceptrons**II-4-1 Définition**

Le mécanisme perceptron fut inventé par le psychologue FRANK Rosenblatt à la fin des années 50. Il représentait sa tentative d'illustrer certaines propriétés fondamentales des systèmes intelligents en général. Le réseau dans ce modèle est formé de trois couches : Une couche d'entrée (la rétine), fournissant des données à une couche intermédiaire, chargée des calculs, cela en fournissant la somme des impulsions qui lui viennent des cellules auxquelles elle est connectée, et elle répond généralement suivant une loi définie avec un seuil, elle-même connectée à la couche de sortie (couche de décision), représentant les exemples à mémoriser. Seule cette dernière couche renvoie des signaux à la couche intermédiaire, jusqu'à ce que leurs connexions se stabilisent.

II-4-2 Architecture d'un perceptron

Le perceptron multicouche est un réseau comportant (L) couches, chaque neurone d'une couche étant totalement connecté aux neurones de la couche suivante.

Chaque neurone (k) est un automate linéaire généralisé dont la fonction de transfert est supposée sigmoïdale.

L'algorithme d'apprentissage par rétro-propagation du gradient de l'erreur est un algorithme itératif qui a pour objectif de trouver le poids des connexions minimisant l'erreur quadratique moyenne commise par le réseau sur l'ensemble d'apprentissage. Cette minimisation par une méthode du gradient conduit à l'algorithme d'apprentissage de rétro-propagation.

Cet algorithme, qui présente l'avantage d'exister, reste discutable dans la mesure où sa convergence n'est pas prouvée. Son utilisation peut conduire à des blocages dans un minimum local de la surface d'erreur. Son efficacité dépend, en effet, d'un grand nombre de paramètres que doit fixer l'utilisateur : le pas du gradient, les paramètres des fonctions sigmoïdes des automates, l'architecture du réseau ; nombre de couches, nombre de neurones par couche..., l'initialisation des poids... [18-19]

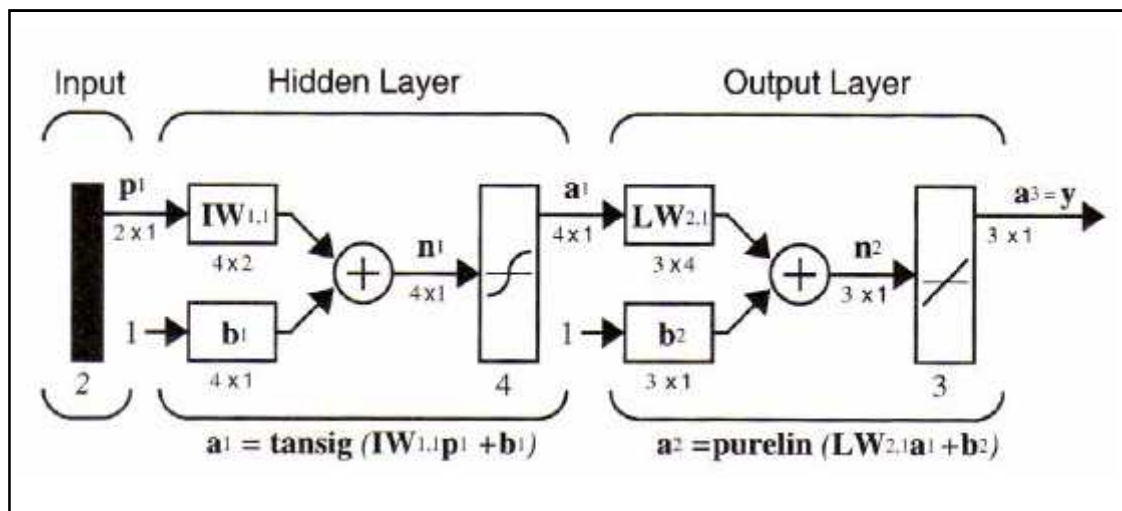


Figure II-16 Perceptron Multicouche

II-4-3 Algorithme d'apprentissage

Les algorithmes d'apprentissage ont pour objectif de permettre à la machine d'acquérir des connaissances opérationnelles à partir d'exemples. Ils cherchent à découvrir la structure cachée d'un ensemble de signaux à partir des valeurs observées, les facteurs inhérents qui expliquent les variations et les dépendances observées. Ils sont utiles pour la prédiction, la prise de décision, et l'analyse de données.

a) Calcul d'activations dans les unités cachées

$$S_j = \sum w_{ji} a_i$$

$$a_j = f(S_j)$$

b) Calcul d'activations dans les unités de sortie

$$S_k = \sum w_{jk} a_j$$

$$a_k = f(S_k)$$

c) Calcul d'erreur entre les sorties désirées et les sorties obtenues

$$e_k = d_k - a_k$$

d) Calcul de l'erreur sur les unités de sortie

$$\delta_k = e_k f'(S_k)$$

e) Calcul de l'erreur sur les unités cachées

$$\delta_j = \sum_k w_{jk} \delta_k f'(S_j)$$

f) Ajustement des poids

$$\text{-Apprentissage des unités cachées : } \Delta W_{ij} = \epsilon \delta_j a_i$$

$$\text{-Apprentissage des unités de sorties : } \Delta W_{jk} = \epsilon \delta_k a_j$$

L'algorithme est schématisé sur la figure (II-17)

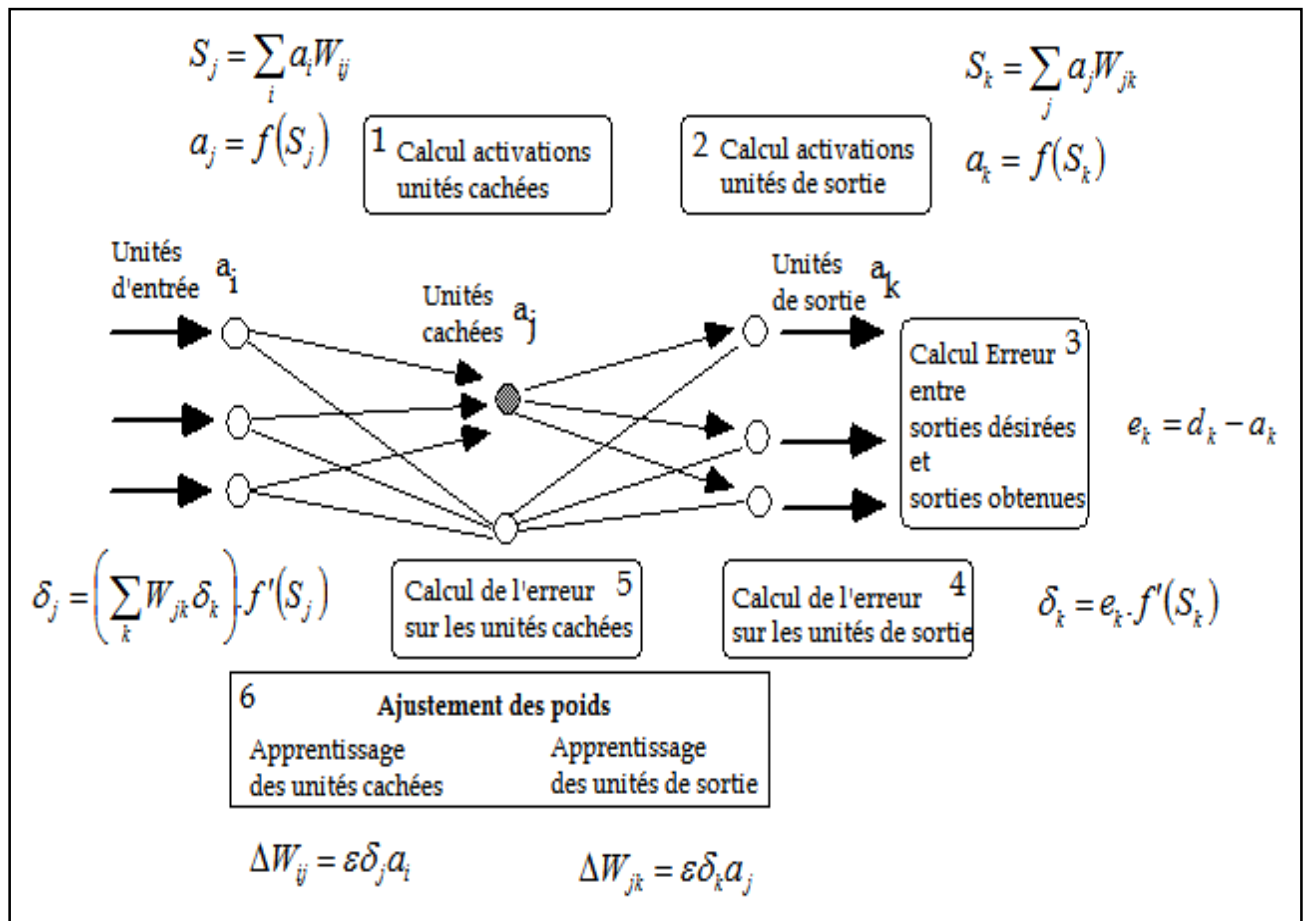


Figure II-17 Algorithme d'apprentissage du perceptron

II-5 Les cartes auto-organisatrices de Kohonen

II-5-1 Définition

Ce type de réseau, appelé aussi SOM (Self Organised Maps) est un réseau à apprentissage non supervisé qui établit une carte discrète, ordonnée topologiquement en fonction de patterns d'entrée. Le réseau forme ainsi une sortie de treillis, où chaque nœud du treillis est un neurone associé à un vecteur de poids. La correspondance entre chaque vecteur de poids est calculée pour chaque entrée. Par la suite, le vecteur de poids ayant la meilleure corrélation, ainsi que certains de ses voisins vont être modifiés afin d'augmenter encore cette corrélation.

Les LVQ (Learning Vector Quantization) sont un cas particulier des SOM, où seul le nœud ayant la meilleure corrélation est adapté. On parle pour ce type de réseau, où seul le 'vainqueur' est sélectionné, de réseaux à compétition.

Une carte auto-organisatrice est un procédé qui convertit un signal d'entrée complexe (plusieurs variables par exemple) en une nouvelle variables catégorielle c'est donc un procédé de classification (modélisation non-supervisée).

Les SOM sont une généralisation de l'analyse en composantes principales.

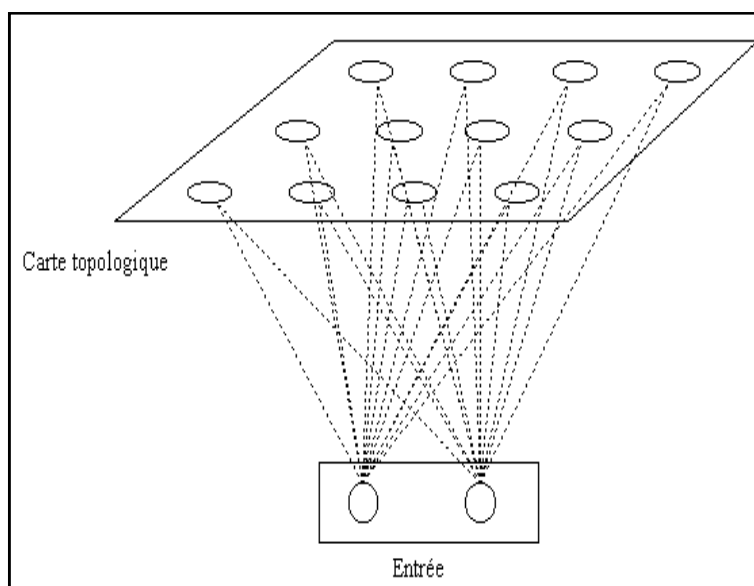
Elle fonctionne comme un réseau de neurones sans variable cible et avec plusieurs nœuds dans la couche de sortie. La carte structure les nœuds en sortie en classes de nœuds. [20]

II-5-2 Architecture

Ces réseaux sont composés d'une grille de neurones (ou nœuds, ou unités) auxquels seront présentés des stimuli. Un stimulus est un vecteur, de dimension d , qui décrit un objet à classer. Ce vecteur peut aussi bien être une description des caractéristiques physiques des objets stimuli qui fait référence à des caractéristiques telle que la présence ou l'absence d'un mot-clé dans un document. Chaque unité de la grille est reliée au vecteur d'entrée (stimulus) par l'intermédiaire de d synapses de poids w (Fig.16). En fait, à chaque unité est associé un vecteur de dimension d qui contient les poids w . C'est la combinaison d'apprentissage non supervisé et d'apprentissage supervisé 3 couches :

- Une couche d'entrée
- Une couche SOM
- Une couche perceptron

La couche auto-organisatrice effectue une sorte de prétraitement des données (réduction de la dimension, regroupement, ...)



**Figure II-18 Schéma d'une carte auto-organisatrice
SOM**

II-5-3 Algorithme d'apprentissage

En premier lieu, la grille doit être initialisée de manière aléatoire.

Un cycle d'apprentissage est constitué des étapes suivantes :

- a) Présenter un vecteur d'entrée associé à un stimulus à la grille.
- b) Trouver le nœud gagnant (ou *winner*). C'est l'unité dont le vecteur associé est le plus similaire au vecteur d'entrée.

$$\| \text{entrée} - \text{neurone}_{winner} \| = \text{Min}_i \| \text{entrée} - \text{neurone}_i \|$$

c) Modifier les poids W_i du nœud gagnant, ainsi que ceux de son entourage, de manière à ce que les vecteurs associés (les vecteurs de poids) «se rapprochent d'avantage» du vecteur d'entrée présenté à la grille. La règle de modification est la suivante :

$$w_i(t+1) = w_i(t) + h(r,t)(p_i - w_i(t)) \quad \text{si } i \text{ voisinage}$$

$$w_i(t+1) = w_i(t) \quad \text{si } i \text{ voisinage}$$

Avec $h(r,t) = \eta \cdot v(r,t)$

η : Le taux d'apprentissage.

$v(r,t)$: la fonction de voisinage)

d) Faire décroître la taille de la zone de voisinage des nœuds gagnants (la zone qui contient les neurones subissant la transformation).

e) Faire décroître le coefficient d'apprentissage, η , qui contrôle l'importance des modifications appliquées aux vecteurs de poids.

f) Arrêter l'apprentissage si le coefficient d'apprentissage est nul, sinon, présenter un autre stimulus à la grille.

La modification des vecteurs associés aux unités se fait de manière différente selon la position des nœuds par rapport à l'unité gagnante. Le nœud gagnant sera celui dont le vecteur subira le plus de modifications, tandis que les unités plus éloignées seront moins affectées. La fonction de voisinage $v(r)$ va être maximale pour $r=0$ et décroître quand r croît (quand on s'éloigne du nœud gagnant).

II-5-4 Domaines d'application

Les cartes de Kohonen permettent le classement d'objets sur lesquels aucune information relative à leur classification n'est disponible a priori. Une fois le réseau organisé (une fois l'apprentissage effectué), chaque objet est classé dans une zone qui correspond à son « gagnant ».

L'utilisation des cartes de Kohonen dans la recherche documentaire consiste à présenter l'ensemble des documents à la carte pour les classer. Ensuite représenter le classement sous forme graphique compréhensible à l'être humain et qui permettra à l'utilisateur de naviguer dans le fonds documentaire pour rechercher l'information désirée.

II-6 Mise en œuvre des réseaux de neurones

Pour réaliser l'approximation de la fonction de régression cherchée, à partir d'échantillons généralement bruités, à l'aide d'un réseau de neurones, trois étapes successives sont nécessaires :

1) Il faut tout d'abord choisir l'architecture du réseau, c'est-à-dire les entrées externes, le nombre de neurones cachés, et l'agencement des neurones entre eux, de telle manière que le réseau soit en mesure de reproduire ce qui est déterministe dans les données. le nombre de poids ajustables est un des facteurs fondamentaux de la réussite d'une application : si le réseau possède un trop grand nombre de poids.

Si le réseau est trop 'souple', il risque de s'ajuster au bruit qui est présent dans les données de l'ensemble d'apprentissage, et, même en l'absence de bruit, il risque de présenter des oscillations non significatives entre les points d'apprentissage, donc de posséder de mauvaises propriétés d'interpolation (ou, dans le jargon des réseaux de neurones, de 'généralisation').

Si par contre ce nombre est trop petit, le réseau est trop 'rigide' et ne peut reproduire la partie déterministe de la fonction. Le problème de la détermination de l'architecture optimale est resté pendant longtemps un problème ouvert, mais il existe actuellement diverses méthodes, mettant notamment en jeu des tests statistiques, qui permettent de déterminer cette architecture pour une vaste classe de réseaux.

2) Il faut calculer ensuite les poids du réseau ou, en d'autres termes, 'estimer les paramètres de la régression non linéaire' à partir des exemples, en minimisant l'erreur d'approximation sur les points de l'ensemble d'apprentissage, de telle manière que le réseau réalise la tâche désirée. Ce calcul des coefficients synaptiques constitue l'apprentissage supervisé pour le réseau de neurones.

3) Il faut estimer enfin la qualité du réseau obtenu, en lui présentant des exemples qui ne font pas partie de l'ensemble d'apprentissage.

II-7 La classification

Supposons que l'on désire classer des formes en deux catégories, A ou B, en fonction de certaines caractéristiques de ces formes ; on peut définir une fonction ϕ qui vaut (+1) pour toutes les formes de la classe A et (-1) pour toutes les formes de la classe B. Les réseaux de neurones sont de bons candidats pour réaliser une approximation de cette fonction ϕ , et on peut démontrer que cette approximation constitue une estimation de la probabilité d'appartenance de la forme inconnue à la classe A.

Les réseaux de neurones fournissent donc une information très riche, qui est loin d'être une simple réponse binaire. Cette propriété remarquable, que les réseaux de neurones partagent avec d'autres classifieurs, n'est malheureusement pas mise à profit dans la plupart des applications.

II-8 Conclusion

Aujourd'hui, on retrouve les réseaux de neurones solidement implantés dans diverses industries, entre autres : dans les milieux financiers, pour la prédiction des fluctuations de marché ; en pharmacie, pour analyser le 'QSAR' (Quantitative Structure-Activity Relationship) ; pour la détection de fraudes sur les cartes de crédit et le calcul de cotes de crédit ; dans les départements de marketing de compagnies de diverses industries, pour prévoir le comportement des consommateurs ; en aéronautique, pour la programmation de pilotes automatiques.

Les applications sont nombreuses et partagent toutes un point commun essentiel à l'utilité des réseaux de neurones : les processus pour lesquels on désire émettre des prédictions comportent de nombreuses variables explicatives et, surtout, il existe éventuellement des dépendances non linéaires de haut niveau entre ces variables qui, si elles sont découvertes et exploitées, peuvent servir à l'amélioration de la prédiction du processus.

CHAPITRE III

Généralités sur les alliages d'aluminium

III-1 Introduction

Parmi l'ensemble des substances solides que l'on rencontre dans la vie quotidienne, il en est certain dont chacun reconnaît le caractère métallique, c'est « *Les métaux* », bien que la plupart du temps il s'agisse en fait d'alliages. La notion de caractère métallique est liée à un certain nombre de propriétés facilement perceptibles comme l'éclat métallique, l'opacité, la bonne conductibilité thermique et électrique, l'élasticité, ou la plasticité. On pourrait, à partir de ces propriétés, élaborer une définition de l'état métallique. Toutefois, cette définition ne serait pas très scientifique et surtout non discriminante, car il est difficile de trouver une propriété qui soit commune à tous les métaux sans être partagée par aucun autre corps non métallique. Il est probable que l'augmentation de leur résistivité électrique avec la température soient les caractéristiques les plus spécifiques aux éléments métalliques.

Il n'existe pas de métal pur entièrement constitué d'un seul type d'atomes, un métal *pur* contient toujours des éléments étrangers, ou impuretés, qui peuvent prendre la forme de défauts ponctuels, il est difficile d'atteindre un degré de pureté dépassant 99,9999. La plus part des métaux n'ont pas ce degrés de puretés ;ce sont des alliages auxquels on ajoute des atomes étrangers pour obtenir des matériaux ayant des caractéristiques précises habituellement les métaux obtenus par alliage ont des meilleurs propriétés mécaniques et ils résistent mieux à la corrosion. [27]

III-2 Les liaisons atomiques

Quand deux atomes se lient, ils peuvent former une molécule. Cette liaison est appelée *liaison primaire*. Toutes les liaisons ont une origine commune : l'attraction électrostatique entre les charges positives du noyau atomique et les charges négatives des électrons. Il y a des attractions et des répulsions entre les charges électriques des électrons et des noyaux. L'équilibre entre cette force de répulsion et la force d'attraction électrostatique conduit à l'existence d'une distance d'équilibre entre deux atomes, caractéristique de la liaison considérée.

Des charges électriques de même signe se repoussent et des charges électriques de signe différent s'attirent.

Le solide est un ensemble d'atomes ou de molécules disposées selon un réseau particulier, la question qu'on peut se poser est de savoir quelles sont les forces qui maintiennent ces atomes et molécules dans un tel état?

Dans un solide la distance inter -atomique ou inter moléculaire est comparable a la taille atomique ou moléculaire .Le solide peut être imaginé comme une molécule géante où les forces de cohésion sont le résultat de l'interaction du nuage électronique et telle que le réseau obtenu forme un système thermodynamiquement stable .cette interaction électronique conduit à plusieurs types de liaisons entre eux la liaison métallique présente dans les métaux et alliages. La liaison métallique concerne des atomes possédant un faible nombre d'électrons sur la couche externe, situation rencontrée dans le cas des métaux. Ces électrons sont faiblement liés au noyau et ils peuvent facilement en être arrachés.

Les atomes sont alors transformés en ions positifs. Les électrons arrachés aux atomes sont mis en commun entre tous ces atomes. Ils constituent un « nuage » ou « gaz » électronique qui assure la cohésion de l'ensemble (Figure III-1), tous ces atomes ont perdu au moins un électron. Comme le cristal est neutre (ils restent dans le cristal, ils ne peuvent pas en sortir sinon le cristal devient positif et attire les e^-). La liaison métallique est donc un ensemble d'ions $+$. Chaque charge positive est entourée par une charge négative. [28]

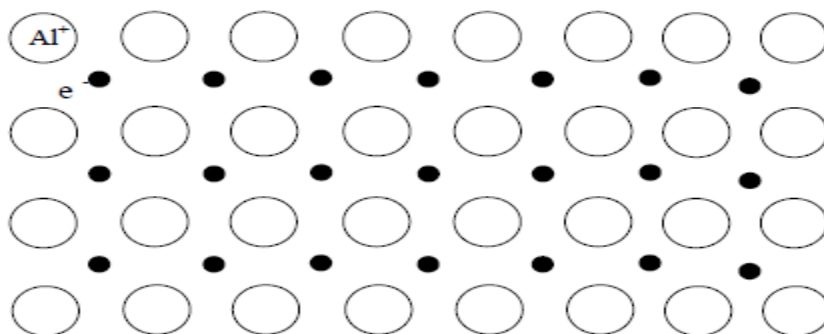


Figure III-1 : Solution solide de substitution

III-3 Structure cristalline

Le **motif** constitue le contenu matériel du cristal, par opposition au réseau, entité purement mathématique qui décrit la géométrie de l'édifice cristallin. Les constituants élémentaires du motif peuvent être de différents types, correspondant à différentes catégories de cristaux. Le motif d'un cristal métallique est donc les atomes contenus dans une maille élémentaire (à spécifier dans chaque cas), le cristal résultant de la répétition tridimensionnelle de ce motif suivant les translations du réseau associé, les métaux se présentent sous trois structures cristallines. [29]

III-3-1 Structure cubique centrée (CC)

Le réseau (sommet du cube), tous ses nœuds sont occupés et il y a un atome au centre et cet atome n'est pas au niveau d'un nœud du réseau. Il faut donc bien distinguer le nœud du réseau qui est une entité mathématique et la maille (contient des atomes) qui est une entité matérielle.

La **maille élémentaire** comporte 2 atomes : Figure III-2

1 au centre du cube et qui donc compte pour 1 et les 8 qui sont au sommet qui comptent pour $1/8$: $8 \times 1/8 = 1 + 1 = 2$.

Exemple de métaux : fer α et chrome

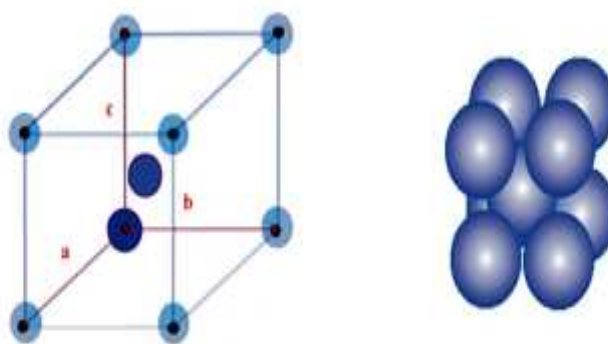


Figure III-2 Structure cubique centrée

III-3-2 Structure cubique à faces centrées(CFC)

La maille élémentaire comporte 4 atomes.

Exemple de métaux : aluminium, cuivre, nickel, fer γ .

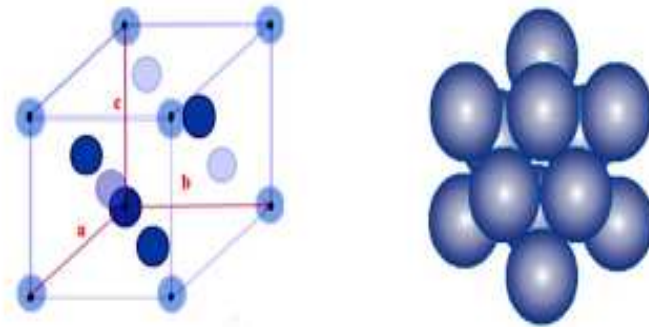


Figure III-3 Structure cubique à faces centrées

III-3-3 Structure hexagonale compacte (HC)

Le réseau associé à cette structure est hexagonal et la maille élémentaire utilisée est le prisme droit à base losange. Le motif contenu dans cette maille est constitué de 2 atomes.

Exemple de métaux : zinc, magnésium, zirconium α et le titane α .

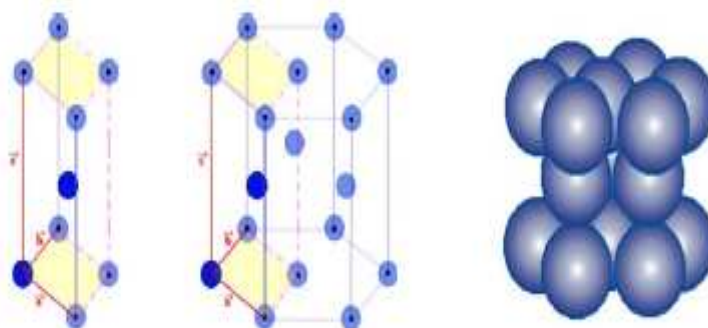


Figure III-4 Structure hexagonale compacte

III-4 Les alliages métalliques

Un alliage est constitué d'un nombre de phases solides distinct – une phase est définie comme une part structurellement homogène du système qui est séparée des autres parties. Chaque phase a sa propre structure et ses propriétés associées.

Un alliage est un solide obtenu par cristallisation d'un mélange de au moins 2 métaux liquides. Après solidification, on obtient un solide polycristallin et dans ce réseau, on va trouver des atomes des deux métaux (ex : AuCu).

Il y a essentiellement trois phases différentes qui peuvent former un alliage : un métal pur, une solution solide ou un composé intermétallique.

III-4-1 Formation des solutions solides

Une solution solide est formée lorsque des atomes du soluté suffisamment petits arrivent à se placer entre les atomes de solvant. En pratique, le diamètre des atomes de soluté ne dépasse pas 60% du diamètre des atomes de solvant.

On peut ajouter dans un métal A des atomes étrangers B sans modifier la structure cristalline de l'élément A, la solution solide AB est formée lorsque l'élément B entre en solution dans A. On dit que l'atome B est un constituant de l'alliage.

L'addition d'impuretés dans un corps produit deux types de défauts ponctuels qui engendrent deux types de solutions : les solutions solides de substitution et les solutions solides d'insertion.

III-4-1-1 Solutions solides d'insertion

Une solution solide est formée lorsque des atomes de soluté suffisamment petits arrivent à se placer entre les atomes de solvant. En pratique, le diamètre des atomes de soluté ne dépasse pas 60% du diamètre des atomes de solvant.

Si les atomes de l'élément d'alliage B ont un rayon atomique suffisamment faible, ils vont pouvoir s'introduire dans les interstices de la structure de A, donnant ainsi naissance à une solution solide d'insertion (ou solution solide interstitielle) [30].

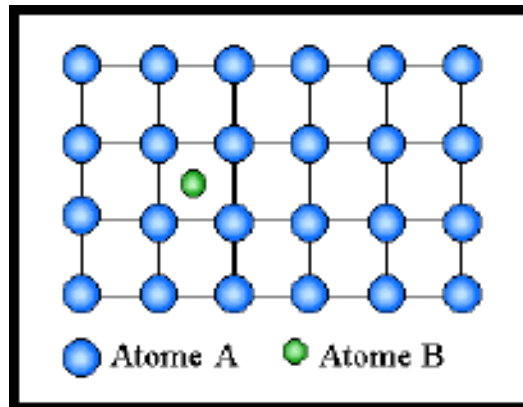


Figure III-5 : Solution solide d'insertion

Il existe deux types d'insertion possibles, l'insertion en *site tétraédrique* où l'atome inséré occupe le centre d'un tétraèdre dont les sommets sont formés par des atomes du solvant métallique, et l'insertion en *site octaédrique* où les atomes insérés occupent le centre d'un octaèdre.

III-4-1-2 Solutions solides de substitution

Ces solutions se forment lorsque les atomes du soluté B occupent des sites normalement occupés par des atomes de A dans le métal pur. L'alliage formé peut être binaire de la forme $(A_{1-x}B_x)$, ternaire de la forme $(A_{1-x-y}B_xC_y)$.

En règle de miscibilité des deux atomes A et B, il faut que ces éléments constituant l'alliage aient la même structure cristalline et des rayons atomiques voisins. [31]

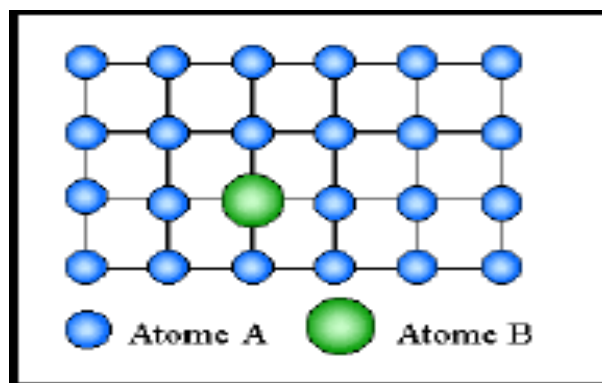


Figure III-6 Solution solide de substitution

III-4-1-3 Règle de Hume-Rothery

La solubilité des atomes en substitution peut être plus ou moins prévue par des lois empiriques, connues sous le nom de **règles de Hume-Rothery** (= règles qui régissent la formation des solutions) :

- Règle des 15% : Pour que B ait une solubilité importante dans A, les rayons de ces éléments ne doivent pas différer de plus de 15%
- Différence d'électronégativité : Si l'un des éléments est plus électropositif par rapport à l'autre, il y a tendance à la formation de liaison AB .plus l'électronégativité est grande entre A et B plus il y a formation d'un composé intermétallique au lieu d'une solution solide.
- Règle de valence : la solubilité est plus grande pour les éléments de valence plus élevée dans un solvant de valence plus faible.
- Structure cristalline : pour que le métal A puissent dissoudre une quantité appréciable de soluté (métal B), la structure cristalline des deux constituants doit être la même.

III-5 Les alliages d'aluminium

Les alliages d'aluminium sont largement utilisés dans l'industrie aéronautique du fait de leur légèreté et de leurs bonnes propriétés mécaniques. Ils sont néanmoins sensibles, dans certains milieux, à différentes formes de corrosion localisée.

Ainsi, pour comprendre la sensibilité à la corrosion il apparaît donc nécessaire de mieux comprendre le comportement des différentes phases métalliques présentes dans ce type de matériau.

L'aluminium pur, non allié possède des propriétés mécaniques faibles, afin d'améliorer ces propriétés, on est conduit à ajouter les éléments d'addition lors de sa fusion, tels que Mg, Mn, Cu, Si, Ni, Ti, Zn, Co, etc. Ces éléments entrent en solution solide en formant des précipités qui entraînent des compositions d'alliages industriels. L'avantage des alliages d'aluminium est qu'ils permettent d'avoir une bonne résistance mécanique tout en conservant une faible masse volumique.

a. Addition de Mg : Les alliages formés, possèdent une résistance mécanique moyenne. Ils sont utilisés dans le domaine de placage, lorsqu'avec l'aluminium, il s'avère difficile. Ils sont faciles à polir et conservent longtemps leur poli à cause de la haute résistance chimique.

b. Addition de (Mg + Mn) : Ces deux éléments augmentent la résistance mécanique sans modifier la déformabilité. Ils se prêtent bien à l'emboutissage et au polissage.

c. Addition (le (Cu + Si) : Ce sont des alliages de fonderie, présentant une bonne coulabilité et conviennent surtout pour les pièces compliquées.

d. Addition de (Co + Ti + Zn) : Accroissent la résistance mécanique sans altérer la coulabilité, l'addition de Cu et de Zn facilite l'usinage.

e. Addition de (Mg + Zn + Cu) : Ils sont aptes aux traitements mécaniques. L'aluminium pur, non allié possède des propriétés mécaniques faibles, afin d'améliorer ces propriétés, on est conduit à ajouter les éléments d'addition lors de sa fusion, tels que Mg, Mn, Cu, Si, Ni, Ti, Zn, Co, etc. Ces éléments entrent en solution solide en formant des précipités qui entraînent des compositions d'alliages industriels. L'avantage des alliages d'aluminium est qu'ils permettent d'avoir une bonne résistance mécanique tout en conservant une faible masse volumique.

III-5-1 L'aluminium

L'aluminium est un élément du IIIème groupe de la classification de Mendeleïev, dont le nombre atomique est 13 et la masse atomique 26,98 (environ 27), possédant une structure cristalline de type CFC (cubique à faces centrées).

L'aluminium représente environ 7,5 % en masse dans l'écorce terrestre, il doit sa place à un ensemble de propriétés qui en font un matériau irremplaçable, et le plus utilisés après le fer et l'acier.

L'aluminium présente une structure cristalline de type cubique à faces centrées (CFC) représentée par le figure (III-4), de paramètre de maille 4.0412\AA . Cette structure est la plus favorable à l'insertion par rapport aux deux autres structures qui sont cubiques centrées(CC) et hexagonal compacte(HC). En effet la distance d'insertion dans la structure CFC est de l'ordre de $0.41d_{\text{solvant}}$, alors que dans CC et HC sont respectivement $0.15 d_{\text{solvant}}$ et $0.225 d_{\text{solvant}}$

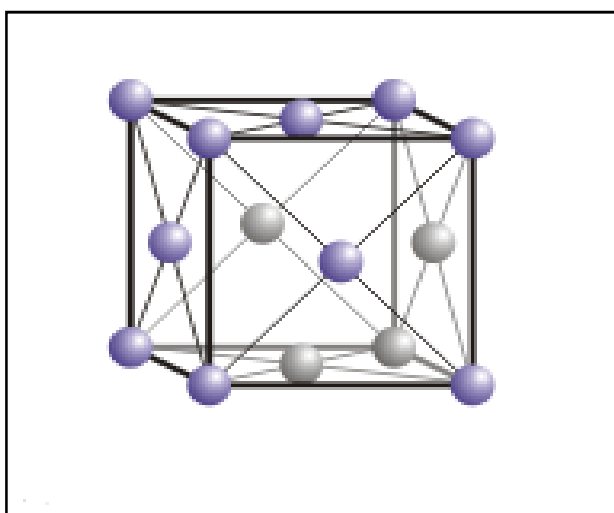


Figure III-7 : Structure cubique à faces centrées

III-5-1-1 Caractéristiques physiques

a) La masse volumique :

C'est l'une des principales caractéristiques, d'une valeur de 2.7 g/cm^3 près de trois fois moins que l'acier, ce qui lui impose une utilisation en constructions aéronautiques et de l'industrie de transport, pour réaliser des économies d'énergie.

b) Résistance à la corrosion

Le phénomène de la tenue à la corrosion est expliqué par la formation d'une pellicule ($\approx 0,7 \mu\text{m}$) et très résistante de Al_2O_3 non soluble dans l'eau et protégeant le métal contre les milieux oxydants et une éventuelle attaque chimique. L'aluminium est très résistant contre l'influence atmosphérique et contre beaucoup de produits de l'industrie chimique et alimentaire (plus l'aluminium est pur et plus sa tenue en corrosion est élevée).

c) Conductivité électrique

Elle représente 62% du Cu pour l'aluminium pur à 99.5%, par ailleurs la masse volumique du cuivre est plus élevée que celle de l'aluminium, par conséquent à caractéristiques électriques égales, et pour une même intensité de courant, un conducteur en aluminium s'échauffe moins.

III-5-1-2 Propriétés mécaniques

L'aluminium pur est très bien déformable à froid et à chaud, il est donc très ductile. Cette plasticité élevée à l'état recuit rend très aisé le corroyage, par contre, son usinage étant mal aisé. L'aluminium se prête bien à toutes les modalités de soudage sous gaz inerte tels que l'argon ou l'hélium (gaz de protection), sinon l'aluminium s'oxyde facilement et forme une pellicule de Al_2O_3 .

La dureté est définie comme étant la résistance à la pénétration locale du matériau, l'aluminium présente une faible dureté. Sa faible dureté et sa faible limite élastique sont défavorables pour son emploi en mécanique. Les propriétés mécaniques de l'aluminium dépendent :

- Des traitements mécaniques (forgeage, laminage, etc.).
- Des traitements thermiques (trempe, recuit, etc.).
- Des éléments d'addition (Fe, Cu, Mn, Zn, Mg, Si).

Une déformation plastique à froid de l'aluminium industriel augmente sa charge de rupture jusqu'à 150 N/mm² mais l'allongement baisse jusqu'à 6 %.

L'aluminium et ses alliages ont une structure cubique à face centrées, ils font preuve à l'état recuit, d'une excellence ductilité. L'aluminium a une température de fusion relativement basse (660°C) et un point d'ébullition de 2060°C, il en résulte une facilité de fusion qui présente un avantage pour fonderie. L'aluminium étant très ductile, on peut facilement le mettre en forme à l'état solide, par déformation plastique (laminage, filage, étirage) soit à chaud soit à température ambiante. [32]

	Fe	Cu	Al	Ni	Mg	Ti	Mo	W
Module de Young (GPa)	211	130	70	200	45	120	325	410
Coéf.de Poisson	0.293	0.343	0.345	0.312	0.291	0.361	0.293	0.28

Tableau III-1 Module d'élasticité et coefficient de Poisson de quelques métaux

III-5-2 Les différents types d'alliage d'aluminium

Il existe deux grandes familles d'alliages, les alliages pour fonderies et les alliages corroyés.

Alliages corroyés : destinés au forge et le laminage, ils sont prévus pour l'obtention des demi-produits (tôles, plaques, feuillards, barres. Profilés, tubes, etc.), ainsi que des pièces de forge et estampées produites par laminage, pressage, forgeage et estampage.

Alliages de fonderies : prévus pour le moulage des formes.

Les différents types sont désignés par un nombre de quatre chiffres et sont classés en fonction des éléments d'alliages comme indiqué dans le tableau III-2

Série	Désignation	Elément d'alliage principal	Phase principal présente dans l'alliage
	1xxx	99%d'aluminium au minimum	
2000	2xxx	Cuivre (Cu)	Al ₂ Cu – Al ₂ CuMg
3000	3xxx	Manganèse(Mn)	Al ₆ Mn
4000	4xxx	Silicium(Si)	
5000	5xxx	Magnésium(Mg)	Al ₃ Mg ₂
6000	6xxx	Magnésium(Mg) +Silicium(Si)	Mg ₂ Si
7000	7xxx	Zinc(Zn)	MgZn ₂
8000	8xxx	Lithium(Li) et autres	Al ₃ Li

Tableau III-2 des différents alliages d'aluminium corroyés

Série	Elément d'alliage principal
100.0	Al commercial
200.0	Al-Cu
300.0	Al-Si-Mg et Al-Si-Cu
400.0	Al-Si
500.0	Al-Mg
700.0	Al-Zn
800.0	Al-Sn

Tableau III-3 des différents alliages d'aluminium de fonderie

III-5-3 Alliages sans durcissement structural

Parmi les huit classes illustrées dans le tableau II-2, une autre distinction peut être faite entre les alliages non trempant correspondent aux séries 1000, 3000 et 5000 (alliages sans durcissement structural) et les alliages trempant à durcissement structural, aux séries 2000, 6000 et 7000. [39]

III-5-3-1 La série 1000 (alliage non allié)

Cette série d'alliages ne peut être traitée thermiquement, la résistance à la traction varie entre 69 à 186 N/mm² elle a une grande résistance à la traction dans un large domaine de température, cette série est souvent désignée comme aluminium pur (min 99.99%). Ces alliages présentent un faible intervalle de fusion et comme ils ont de faibles propriétés mécaniques ils sont rarement utilisés en construction.

Les alliages de la série (1000) sont utilisés en raison de leur résistance élevée à la corrosion et leur bonne conductibilité électrique.

III-5-3-2 La série 3000 (alliage Al-Mn)

Ces alliages sont à base d'aluminium et de manganèse la résistance à la traction est de 110 à 283 N/mm² a teneur en manganèse qui varie entre .La résistance de ces alliages est modérée mais ils possèdent une bonne résistance à la corrosion, et pouvant être utilisés à haute température.

III-5-3-3 Série 5000 (Al-Mg)

A ce groupe appartiennent les alliages avec une teneur en magnésium de 0,2 à 7,2 % et une teneur en Mn de 0,2 à 0,8 %. Ces alliages sont caractérisés par une résistance (stabilité) élevée contre l'eau salée. Les alliages avec une teneur en Mn jusqu'à 5 % s'adaptent bien à l'oxydation anodique à cause de leur structure fine et homogène.

Les alliages avec une teneur ($> 5\%$) en Mn se laissent difficilement travailler et posent des problèmes aux soudages. Les alliages à 7% Mn sont caractérisés par de bonnes propriétés de surface.

A l'état d'équilibre, après refroidissement, les alliages Al-Mg sont biphasés $\alpha + \beta$, (Al_3Mg_2). Pourtant la haute stabilité de la solution solide et la faible vitesse de diffusion du Manganèse dans l'aluminium font que même, après refroidissement lent, ils ne contiennent pas de phase excédentaire et se composent d'équilibre partiel.

III-5-4 Alliages à durcissement structural

Le durcissement des alliages d'aluminium a été découvert en 1906 par l'allemand Alfred Wilhr dans le système Al-Cu-Mg. Il est très connu aujourd'hui pour plusieurs alliages. Le but de ce traitement est d'augmenter la résistance mécanique et la résistance à l'usure du matériau. Les différents alliages d'aluminium destinés au durcissement sont : AlCuMg, AlMgSi, AlZnMg, AlZnMgCu.

III-5-4-1 La série 2000 (alliage Al-Cu)

L'élément d'addition est le cuivre avec une teneur qui varie entre 0.7 à 6.8% et celle du magnésium entre 0.5 à 1.5% . Ces alliages peuvent être traitée thermiquement, la résistance à la traction varie entre 186 à 427 N/m^2 donc à résistance très élevée par rapport à la série précédente.

Ces alliages sont utilisés dans l'industrie aéronautique et spatiale en raison de leur durcissement structural qui leur confère d'excellentes propriétés mécaniques.

III-5-4-2 La série 4000 (alliage Al-Si)

Ces alliages sont en aluminium silicium d'une teneur entre 0.6 à 21.5% en Si, c'est la seule série susceptible de traitement thermique. Les propriétés mécaniques de cette structure sont faibles.

III-5-4-3 Série 6000 (Al-Mg/Si)

Ces alliages ont une très bonne aptitude à la transformation à chaud, ils sont très résistants à la corrosion. Leur durcissement est dû à la précipitation de la phase durcissante Mg_2Si . Ce groupe d'alliage comprend les alliages avec (0,4 à 1,4 %) Mg, (0,3 à 1,6 %) Si et de (0 à 1 %) Mn. On remarque qu'un excès de silicium dans cette série d'alliages améliore les propriétés mécaniques mais au détriment de la corrosion. L'addition de cuivre améliore également les propriétés mécaniques mais la teneur est limitée à 05% la aussi a cause de la résistance à la corrosion.

III-5-4-4 Série 7000 (Al-Zn-Mg et Al-Zn-Mg/Cu)

Les alliages appartenant à ce groupe ont une teneur en Zn de (4 à 5 %), de Mg (0,5 à 3,5 %), de Mn (0 à 1 %), et de Cr (0 à 0,3 %). L'alliage avec (4 à 5 %) Zn, (0,5 à 1,5 %) Mg, (0 à 1 %) Mn et (0 à 0,3%) Cr se durcit à la température de 350 à 500°C, possédant de bonnes propriétés de déformation .

Avec l'augmentation de la teneur en Zn et Mg, la résistance de l'alliage croît, alors que la plasticité et la tenue à la corrosion diminuent. La tenue à la corrosion et les propriétés mécaniques peuvent être améliorés par l'addition de faibles quantités de manganèse, de chrome et de vanadium ou par un refroidissement lent jusqu'à moyen.

L'alliage Al-Zn-Mg/Cu est une suite développée des alliages Al-Zn-Mg, présentant une haute résistance. L'addition du cuivre et du manganèse augment la résistance mécanique, la bonne tenue à la corrosion est obtenue aussi par l'addition de Mn, Cr, Va. Comme matériaux de plaquage, on utilise un alliage d'aluminium contenant du zinc (1 à 3) %.

L'augmentation de la teneur en Zn et Mg, la résistance de l'alliage croît, alors que leur plasticité et tenue à la corrosion diminuent. Ils ont une bonne plasticité à chaud et se prêtent relativement bien au corroyage à froid après recuit. L'alliage avec (3,8 à 5,2 %) Zn, (2,4 à 3,8 %) Mg, (0,4 à 1,0 %) Cu, (0,1 à 0,4 %) Mn et (0,1 à 0,3 %) Cr est utilisé dans le domaine de construction à sollicitation élevée, aviation, machines, exploitation minière.

Un alliage avec (5,1 à 6,1 %) Zn, (2,1 à 2,9 %) Mg, (1,2 à 2 %) Cu, (0 à 0,3 %) Mn, (0,18 à 0,4 %) Cr, utilisé dans le domaine de l'aviation, le matériel roulant et les machines complexes.

III-6 Le durcissement structural

La résistance et la dureté de certains alliages d'aluminium peuvent être améliorées par une série de traitements thermiques définie par un processus nommé durcissement structural.

Le durcissement de l'aluminium est obtenu par l'ajout d'éléments qui modifient plus ou moins sa microstructure soit en entrant en solution solide de substitution (cas du magnésium), soit en formant des précipités hors solution solide (cas du manganèse), soit en permettant le durcissement structural, ce dernier phénomène, extrêmement important, est le processus par excellence qui permet d'obtenir une résistance mécanique élevée. Il est réalisé par un traitement thermique effectué en trois stades :

- ❖ chauffage à une température suffisamment élevée (450 °C à 550 °C suivant le type d'alliage) pour mettre les éléments d'addition en solution solide (homogénéisation).
- ❖ refroidissement rapide ou trempe pour maintenir la solution solide sursaturée à la température ambiante.
- ❖ vieillissement soit à la température ambiante ou à une température comprise entre 100 °C et 220 °C selon l'alliage.

Ce durcissement structural s'explique par la décomposition en plusieurs étapes de la solution solide sursaturée, les atomes d'éléments d'addition formant de très fins précipités, plus ou moins cohérents avec la matrice d'aluminium.

III-6-1 Durcissement structural dans les alliages d'aluminium

La phase α est une solution solide de substitutions, le cuivre y est dissous dans l'aluminium, pour un alliage à 04% en masse de cuivre.

La température de mise en solution se situe vers 500°C, le domaine de précipitation rapide est voisin de 400°C, à ce stade trois types de précipités peuvent apparaître. Entre 300°C et 400°C un composé intermétallique Al_2Cu de structure quadratique centrée de paramètres $a=0,607$ nm et $c=0,487$ nm (tétraonal, phase θ) incohérent.

Vers 200°C et 300°C, le précipité possède toujours la même structure tétraonal mais sous forme de plaquette θ' avec une orientation parallèle au plan (1 0 0) de la solution solide mère (CFC) . Cette phase a une structure quadratique avec $a=0,404$ nm et $c=0,581$ nm.

A 150°C les précipités apparaissent en forme de disques très fins (diamètre de l'ordre de 150 nm et d'épaisseur 10nm environ) leur structure bien que tétraonal est différente de celle des précipités formés aux températures plus élevées, c'est la phase θ'' . Cette phase se présente sous la forme de plans riches en Cu (toujours dans les plans {100} de la matrice) elle est toujours cohérente et les contraintes élastiques sont importantes. Elle est en partie responsable du maximum de durcissement des alliages du système Al-Cu.

Aux plus basses températures, jusqu'à la température ordinaire, un type de transformation un peu différent se manifeste, les atomes de cuivre se rassemblent sur le plan (1 0 0) sous formes d'amas de plans cohérents avec le réseau appelées amas ou zones de Guinier-Preston ou simplement zones G.P . [33]

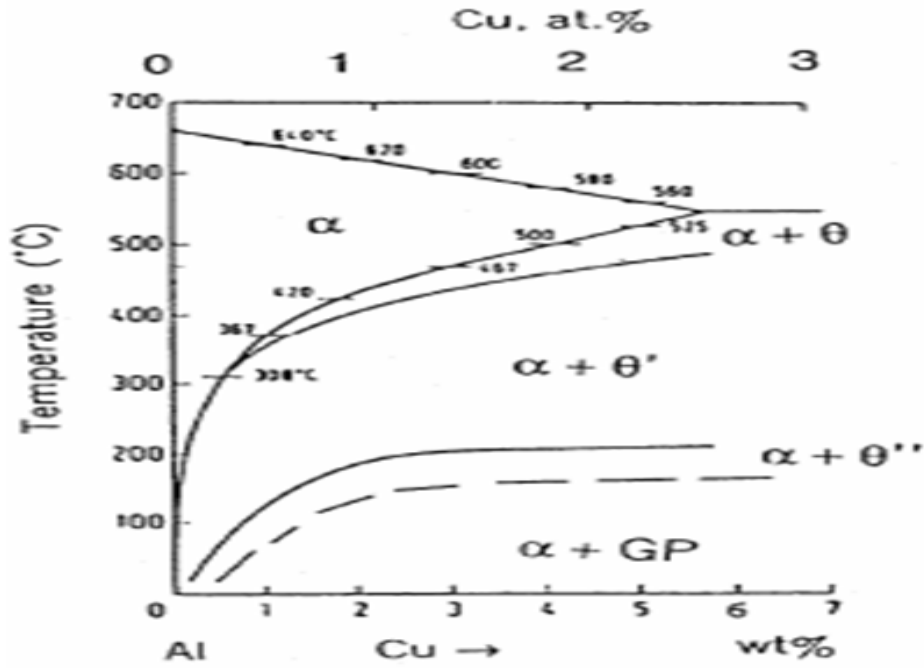


Figure III-8 Une partie d'équilibre du système Al-Cu riches en Al avec les domaines d'apparition des différentes phases

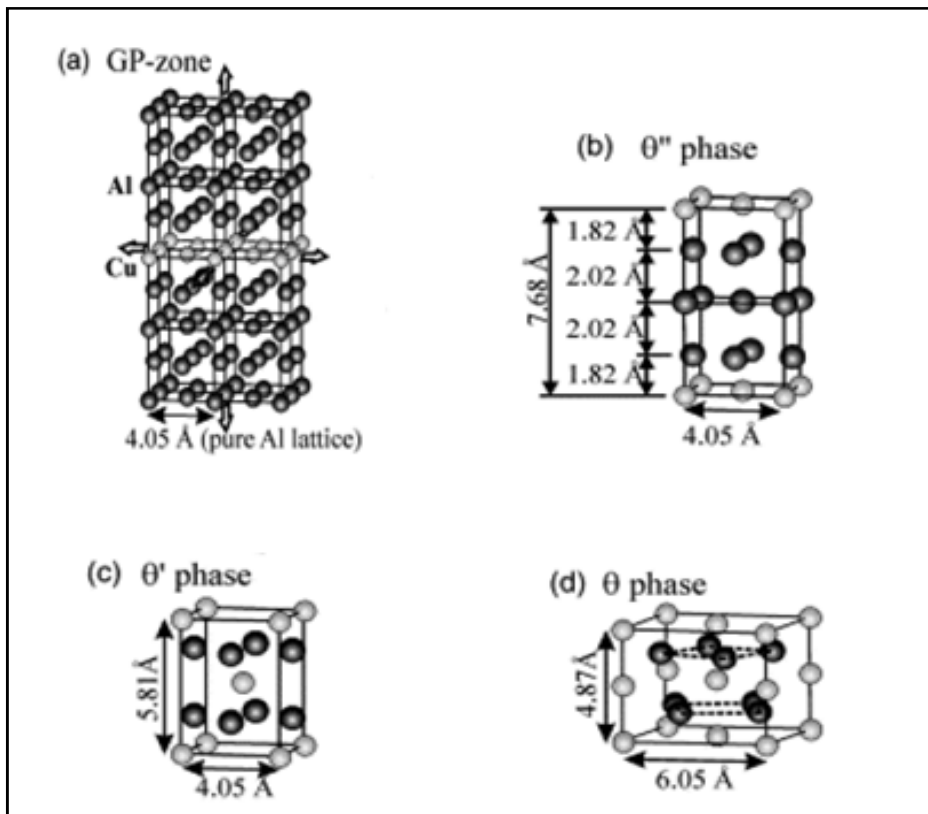


Figure III-9 Structure cristalline des différentes phases

Pour les autres alliages c'est le même procédé sauf que le nombre d'étapes intermédiaires peut être différent de même que la structure comme par exemple la série (7000) d'alliage d'aluminium indiquée sur le tableau (III-4).

Al-Zn-Mg	G.P .zones	G.P .zones sphériques
	η	Hexagonal $MgZn_2$
	η	Hexagonal
	T'	Semi cohérente hexagonal
	T	Incohérente cubic $Mg_{32}(Al, Zn)_{49}$

Tableau III-4 Les différentes phases de l'alliage Al-Zn-Mg

III-6-2 Exemple d'alliages de la série 7000

D'après les tableaux nous remarquons que les alliages d'aluminium de la série 7xxx (Al-Zn/Cu) possèdent des meilleures propriétés mécaniques appelés alliages d'aluminium à haute résistance mécanique, et qui sont le but de notre travail. [34]

III-6-2-1 Série 7005

Mn	Zr	Mg	Cr	Zn	Al	Etat recuit				Etat T6			
						TS MPa	YS MPa	HB	A %	TS MPa	YS MPa	HB	A %
0.45	0.14	1.4	0.13	0.45	Bal	195	80	53	20	350	290	94	13

III-6-2-2 Série 7039

Cu	Mg	Cr	Zn	Al	Etat recuit				Etat T6			
					TS MPa	YS MPa	HB	A%	TS MPa	YS MPa	HB	A%
0.25	2.8	0.2	4.00	Bal	230	100	61	22	400	330	123	13

III-6-2-3 Série 7075

Cu	Mg	Cr	Zn	Al	Etat recuit				Etat T6			
					TS MPa	YS MPa	HB	A%	TS MPa	YS MPa	HB	A%
1.6	2.5	0.23	5.6	Bal	228	103	60	17	572	503	150	11

III-6-2-4 Série 7178

Cu	Mg	Cr	Zn	Al	Etat recuit				Etat T6			
					TS MPa	YS MPa	HB	A%	TS MPa	YS MPa	HB	A%
2.00	2.8	0.26	6.8	Bal	228	103	60	6.8	607	538	160	10

TS : résistance à la traction R_m

YS : limite conventionnelle d'élasticité $R_{p0.02}$

HB : dureté Brinel

A% : élongation

III-6-3 Etat métallurgique des alliages d'aluminium

On peut obtenir des alliages d'aluminium en divers états, recuit, écrouis ou trempés et vieillis. On désigne souvent ces alliages par le différent état, (représenté par une lettre) suivi d'un chiffre.

O	Etat brut de transformation
T1	Refroidi après transformation à chaud et mûri
T2	Refroidi après transformation à chaud, écroui et mûri
T3	Mise en solution, écroui et mûri
T4	Mise en solution et mûri
T5	Refroidi après transformation à chaud puis revenu
T6	Mise en solution puis revenu
T7	Mise en solution puis sur-revenu
T8	Mise en solution, écroui puis revenu
T9	Mise en solution, revenu écroui puis
T10	Ecroui après refroidissement et transformation à chaud et mûri

Tableau -4 représentation des états de traitements de Al-Zn

III-7 Conclusion

Dans ce chapitre on a présenté des généralités sur l'aluminium et ses alliages tout en indiquant comment obtenir un alliage dur et l'influence des éléments d'addition sur les propriétés mécaniques. Parmi les séries d'alliages d'aluminium qui présente de meilleurs propriétés mécaniques la série 7000 dont l'élément d'addition principale est le zinc que nous allons essayer de les classer suivant « dure ductile » et « dure fragile » par l'une des techniques du data mining qui est les réseaux de neurones artificiels. On va appliquer un algorithme d'apprentissage pour une classification supervisée, comme modèle « *le perceptron multicouche* ». A la fin on passe la phase généralisation donc prédiction des propriétés d'un alliage comme le module de compression, le module de cisaillement et le rapport B/G.

CHAPITRE IV :

Application des RNN à l'alliage Al-Zn

IV-1 Introduction

Dans ce chapitre, nous décrivons le modèle de réseau de neurones que nous avons utilisé pour l'analyse du comportement mécanique de l'alliage d'aluminium (Al-Zn) en fonction de la teneur en cuivre.

Un certain nombre de données est nécessaire pour développer un réseau de neurones performant, incluant son architecture, les fonctions d'apprentissage, l'algorithme d'apprentissage et les autres paramètres du réseau.

Un processus d'apprentissage et une méthode d'évaluation sont ensuite utilisés. Après avoir entraîné le réseau pour résoudre le problème basé sur ces données, de nouvelles données du même problème peuvent être introduites dans le réseau entraîné pour faire sortir des solutions réelles. Le processus de mise au point du réseau de neurones comporte les quatre étapes suivantes :

1. Collection de données : analyse et traitement de données.
2. Apprentissage du réseau : ceci inclut le choix de son architecture, les fonctions d'apprentissage, l'algorithme d'apprentissage et les paramètres du réseau.
3. Test du réseau entraîné, pour évaluer les performances du réseau.
4. Utilisation du réseau entraîné pour la simulation et la prévision

IV-2 La base de données

Alliages (Al-Zn)	Al	Cr	Cu	Fe	Mg	Mn	Si	Ti	Zn	dureté	densité	A%	B GPa	G GPa	B/G
7001 O	0,87	0,35	2,60	0,40	3,40	0,20	0,35	0,20	8,00	60	2,84	14	69,6078	27,00	2,5781
7001 T6	0,87	0,35	2,60	0,40	3,40	0,20	0,35	0,20	8,00	160	2,84	9	69,6078	27,00	2,5781
7001 T7	0,87	0,35	2,60	0,40	3,40	0,20	0,35	0,20	8,00	155	2,84	12	69,6078	27,00	2,5781
7005 T6	91,00 94,70	0,20	0,10	0,40	3,40	0,70	0,35	0,00	5,00	94	2,78	13	70,5882	26,90	2,6241
7005 T53	91,00 94,70	0,20	0,10	0,40	1,80	0,70	0,35	0,00	5,00	105	2,78	15	70,5882	26,90	2,6241
7039 T61	90,45	0,25	0,10	0,40	3,30	0,00	0,30	0,10	4,50	123	2,74	13	68,2353	26,00	2,6244
7039 T64	90,45	0,25	0,10	0,40	3,30	0,00	0,30	0,10	4,50	133	2,74	13	68,2353	26,00	2,6244
7050 T74	87,3 90,30	0,04	2,60	0,15	2,60	0,10	0,12	0,06	6,70	140	2,83	11	70,2941	26,90	2,6132
7050 T76	87,3 90,30	0,04	2,60	0,15	2,60	0,10	0,12	0,06	6,70	147	2,83	11	70,2941	26,90	2,6132
7050 T73	87,3 90,30	0,04	2,60	0,15	2,60	0,10	0,12	0,06	6,70	132	2,83	12	70,2941	26,90	2,6132
7075 O	87,1 91,40	0,28	2,00	0,50	2,90	0,30	0,40	0,20	6,10	60	2,81	17	70,2941	26,90	2,6132
7075 T6	87,1 91,40	0,28	2,00	0,50	2,90	0,30	0,40	0,20	6,10	150	2,81	11	70,2941	26,90	2,6132
7075 T73	87,1 91,40	0,28	2,00	0,50	2,90	0,30	0,40	0,20	6,10	135	2,81	13	70,5882	27,00	2,6144
7175 T66	88 91,40	0,28	2,00	0,20	2,90	0,10	0,15	0,10	6,10	150	2,80	11	70,5882	27,00	2,6144
7175 T736	88 91,40	0,28	2,00	0,20	2,90	0,10	0,15	0,10	6,10	145	2,80	12	70,5882	27,00	2,6144
7175 T7351	88 91,40	0,28	2,00	0,20	2,90	0,10	0,15	0,10	6,10	135	2,80	13	70,5882	27,00	2,6144
7475 T61	88,5 91,50	0,25	1,90	0,12	2,60	0,06	0,10	0,06	6,20	148	2,81	11	68,9216	27,00	2,5527
7475 T65	88,5 91,50	0,25	1,90	0,12	2,60	0,06	0,10	0,06	6,20	150	2,81	13	70,2941	27,00	2,6035
7475 T761	88,5 91,50	0,25	1,90	0,12	2,60	0,06	0,10	0,06	6,20	140	2,81	12	68,9216	27,00	2,5527
7475 T7351	88,5 91,50	0,25	1,90	0,12	2,60	0,06	0,10	0,06	6,20	135	2,81	13	70,2941	27,00	2,6035
7475 T7651	88,5 91,50	0,25	1,90	0,12	2,60	0,06	0,10	0,06	6,20	140	2,81	12	70,2941	27,00	2,6035

Référence [43]

Tableau IV-1 : Bases de données

IV-3 Le réseau de neurones

Le réseau de neurones est un réseau avec des nœuds ou neurones analogues aux neurones biologiques. Les nœuds sont interconnectés à des liens de pondération.

Les poids sont ajustables et peuvent être entraînés par un processus et des exemples d'apprentissage. Il y a différentes couches interconnectées dans un réseau complexe. La figure IV-1 montre la structure d'un réseau de neurones avec plusieurs couches.

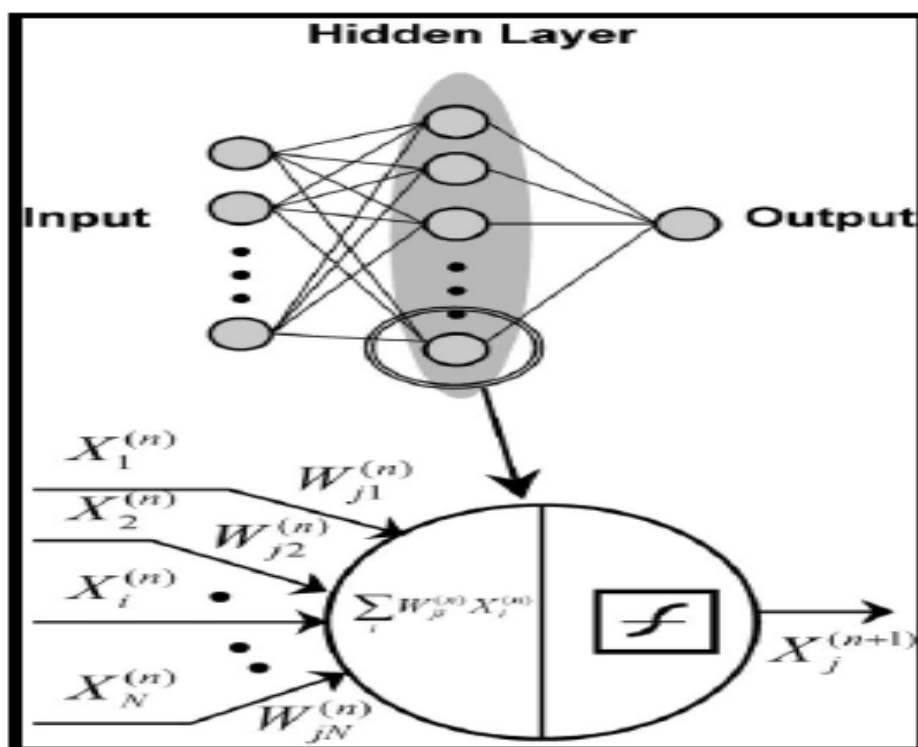


Figure IV-1 : Description schématique d'un réseau de neurones

Nous avons utilisé un réseau de neurones de type perceptron multicouches (MLP) avec un algorithme d'apprentissage de type rétro propagation. Les concentrations de l'aluminium et des différents éléments d'addition ainsi que les six propriétés qui sont la dureté(HB), la densité, l'allongement (A%), le module de compressibilité(B), le module de cisaillement (G) et le ratio B/G sont présentées comme entrées du réseau. La propriété d'être « dure et ductile » (avec une teneur en cuivre inférieure à 2%), et « dure et fragile » (avec une teneur en cuivre supérieure à 2%) correspondent à la sortie du réseau.

Dans le processus d'apprentissage, un ensemble de données, regroupant les entrées et les sorties correspondantes a été utilisé pour entraîner le réseau. Un algorithme itératif ajuste les poids, de telle sorte que les sorties du réseau (Eq IV.1) soient les plus proches possibles des sorties désirées.

Cet objectif est atteint en minimisant l'erreur d'apprentissage, définie par l'erreur quadratique moyenne (Eq IV.2)

$$y_i = f \left(\sum_{j=1}^n w_{ij} x_j + b_i \right) \quad (\text{Eq IV.1})$$

Où :

- y_i est la sortie, x_i est l'entrée

- f est la fonction d'activation, c'est une fonction non linéaire d'une combinaison des entrées pondérées par les paramètres w_{ij} .

- w_{ij} sont les poids, ce sont des paramètres de pondération ajustables de la fonction d'activation.

- b_i est le biais, c'est un paramètre constant de correction

$$\text{MSE} = \frac{1}{QN} \sum_{m=1}^Q \sum_{n=1}^N (d_n(m) - y_n(m))^2 \quad (\text{Eq IV.2})$$

Où :

d est la sortie désirée,

y est la sortie du réseau

Q est le nombre d'ensembles d'apprentissage

N est le nombre de sorties.

IV-4 L'architecture du réseau

Le réseau que nous avons utilisé présente les caractéristiques suivantes :

- Une couche d'entrée : dans laquelle nous présentons en entrées 21 échantillons sous forme de vecteurs de taille 14 (14 paramètres), qui sont les concentrations des éléments d'addition et les six propriétés mécaniques. Le nombre de nœuds (neurones) en entrées est donc égale au nombre de colonnes c'est-à-dire au nombre de paramètres pour chaque alliage.
- Une couche de sortie, dans laquelle nous présentons les sorties, qui sont les classes présentées par un code binaire 0 ou 1. Dans notre cas, on a deux classes, la première classe est « dure et ductile » pour une teneur en cuivre inférieure à 2% codée par 0. La deuxième classe est « dure et fragile » pour une teneur en cuivre supérieure à 2% codée par 1. En sortie donc un seul nœud (neurone).
- Apprentissage du réseau : pour faire entraîner le réseau, nous avons utilisé un algorithme itératif de rétro-propagation ; avec vingt et un échantillons (alliages de la série 7000) avec différentes concentration en aluminium et en éléments d'addition et les six propriétés avec comme exemples d'apprentissage.
- Une seule couche cachée : constituée de 15 neurones.
- La fonction d'activation pour la couche cachée est une logarithmique sigmoïde, elle a la forme suivante :

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

IV-5 Application à l'étude de l'alliage (Al-Zn)

IV-5-1 Apprentissage du réseau

Pour l'apprentissage du comportement élastique des alliages d'aluminium (Al-Zn), nous avons utilisé une base de données (BS1) qui représente 21 échantillons comme indiqué dans le tableau IV-2.

En entrée, les 14 paramètres qui sont les différentes concentrations des 21 alliages ainsi que les six propriétés mécaniques qui sont la dureté, la densité, l'allongement, le module de compression B, le module de cisaillement G et le ratio B/G. En sortie (out1) dans laquelle on a opté pour un codage binaire, 0 pour « dure et ductile », et 1 pour « dure et fragile », avec un réseau de neurones artificiels (net1) définis par la fonction « newff ». Les données ont été toutes normalisées à 100.

La teneur en cuivre est prise en considération en effet les deux classes binaires représentent respectivement 0 pour une teneur inférieure à 2% en cuivre et 1 pour une teneur supérieure à 2%.

La sortie supervisée (out1) est représentée sur le tableau IV-2.

														Out 1
0,87	0,35	0,40	3,40	0,20	0,35	0,20	8,00	60	2,84	14	69,6078	27,00	2,5781	1
0,87	0,35	0,40	3,40	0,20	0,35	0,20	8,00	160	2,84	9	69,6078	27,00	2,5781	1
0,87	0,35	0,40	3,40	0,20	0,35	0,20	8,00	155	2,84	12	69,6078	27,00	2,5781	1
91,00	0,20	0,40	3,40	0,70	0,35	0,00	5,00	94	2,78	13	70,5882	26,90	2,6241	0
91,00	0,20	0,40	1,80	0,70	0,35	0,00	5,00	105	2,78	15	70,5882	26,90	2,6241	0
90,45	0,25	0,40	3,30	0,00	0,30	0,10	4,50	123	2,74	13	68,2353	26,00	2,6244	0
90,45	0,25	0,40	3,30	0,00	0,30	0,10	4,50	133	2,74	13	68,2353	26,00	2,6244	0
87,3	0,04	0,15	2,60	0,10	0,12	0,06	6,70	140	2,83	11	70,2941	26,90	2,6132	1
87,3	0,04	0,15	2,60	0,10	0,12	0,06	6,70	147	2,83	11	70,2941	26,90	2,6132	1
87,3	0,04	0,15	2,60	0,10	0,12	0,06	6,70	132	2,83	12	70,2941	26,90	2,6132	1
87,1	0,28	0,50	2,90	0,30	0,40	0,20	6,10	60	2,81	17	70,2941	26,90	2,6132	1
87,1	0,28	0,50	2,90	0,30	0,40	0,20	6,10	150	2,81	11	70,2941	26,90	2,6132	1
87,1	0,28	0,50	2,90	0,30	0,40	0,20	6,10	1,35	2,81	13	70,5882	27,00	2,6144	1
88	0,28	0,20	2,90	0,10	0,15	0,10	6,10	1,50	2,80	11	70,5882	27,00	2,6144	1
88	0,28	0,20	2,90	0,10	0,15	0,10	6,10	1,45	2,80	12	70,5882	27,00	2,6144	1
88	0,28	0,20	2,90	0,10	0,15	0,10	6,10	1,35	2,80	13	70,5882	27,00	2,6144	1
88,5	0,25	0,12	2,60	0,06	0,10	0,06	6,20	1,48	2,81	11	68,9216	27,00	2,5527	0
88,5	0,25	0,12	2,60	0,06	0,10	0,06	6,20	1,50	2,81	13	70,2941	27,00	2,6035	0
88,5	0,25	0,12	2,60	0,06	0,10	0,06	6,20	1,40	2,81	12	68,9216	27,00	2,5527	0
88,5	0,25	0,12	2,60	0,06	0,10	0,06	6,20	1,40	2,81	13	70,2941	27,00	2,6035	0
88,5	0,25	0,12	2,60	0,06	0,10	0,06	6,20	1,40	2,81	12	70,2941	27,00	2,6035	0

Tableau IV-2 : La base d'apprentissage (BS1) des données entrées avec les sorties (out1)

Le programme implémenté sous MATLAB donne la courbe représentée sur la figure IV-2 qui montre à quel point le net a bien été entraîné.

En effet l'erreur quadratique dans la phase apprentissage du réseau de la fonction de performance est de 0.09866 (SSE=0.0986659) pour un objectif de 0.1 (Goal=0.1) donc assez bonnes performances.

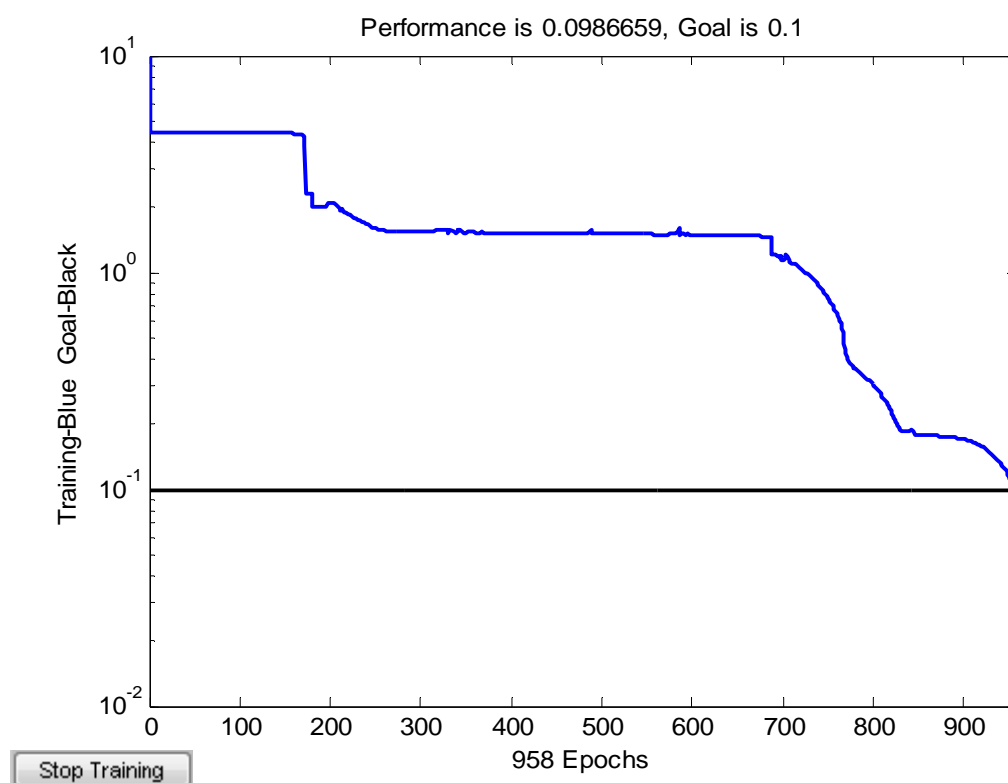


Figure IV-2 : Performance du réseau dans la phase apprentissage

IV-5-2 Présentation de la deuxième base d'exemples au réseau

Toujours avec la même architecture du réseau de neurones mais on va l'appeler (net2) et on l'entraîne avec un 2^{ème} ensemble de données avec 18 alliages d'aluminium cette fois ci, en gardant les mêmes concentrations des éléments d'additions et les six propriétés élastiques c'est-à-dire les 14 paramètres, avec une variation de la concentration en aluminium. La base de données (BS2) en entrée et (out2) en sortie (la sortie désirée) est représentée sur le tableau IV-3.

														Out 2
94,7	0,20	0,40	3,40	0,70	0,35	0,00	5,00	0,94	2,78	13	70,5882	26,9	2,6241	0
94,7	0,20	0,40	1,80	0,70	0,35	0,00	5,00	1,05	2,78	15	70,5882	26,9	2,6241	0
90,5	0,25	0,40	3,30	0,00	0,30	0,10	4,50	1,23	2,74	13	68,2353	26,00	2,6244	0
90,5	0,25	0,40	3,30	0,00	0,30	0,10	4,50	1,33	2,74	13	68,2353	26,00	2,6244	0
90,3	0,04	0,15	2,60	0,10	0,12	0,06	6,70	1,40	2,83	11	70,2941	26,9	2,6132	1
90,3	0,04	0,15	2,60	0,10	0,12	0,06	6,70	1,47	2,83	11	70,2941	26,9	2,6132	1
90,3	0,04	0,15	2,60	0,10	0,12	0,06	6,70	1,32	2,83	12	70,2941	26,9	2,6132	1
91,4	0,28	0,50	2,90	0,30	0,40	0,20	6,10	0,60	2,81	17	70,2941	26,9	2,6132	1
91,4	0,28	0,50	2,90	0,30	0,40	0,20	6,10	1,50	2,81	11	70,2941	26,9	2,6132	1
91,4	0,28	0,50	2,90	0,30	0,40	0,20	6,10	1,35	2,81	13	70,5882	27,00	2,6144	1
91,4	0,28	0,20	2,90	0,10	0,15	0,10	6,10	1,50	2,80	11	70,5882	27,00	2,6144	1
91,4	0,28	0,20	2,90	0,10	0,15	0,10	6,10	1,45	2,80	12	70,5882	27,00	2,6144	1
91,4	0,28	0,20	2,90	0,10	0,15	0,10	6,10	1,35	2,80	13	70,5882	27,00	2,6144	1
91,5	0,25	0,12	2,60	0,06	0,10	0,06	6,20	1,48	2,81	11	68,9216	27,00	2,5527	0
91,5	0,25	0,12	2,60	0,06	0,10	0,06	6,20	1,50	2,81	13	70,2941	27,00	2,6035	0
91,5	0,25	0,12	2,60	0,06	0,10	0,06	6,20	1,40	2,81	12	68,9216	27,00	2,5527	0
91,5	0,25	0,12	2,60	0,06	0,10	0,06	6,20	1,35	2,81	13	70,2941	27,00	2,6035	0
91,5	0,25	0,12	2,60	0,06	0,10	0,06	6,20	1,40	2,81	12	70,2941	27,00	2,6035	0

Tableau IV -3 : Deuxième base d'exemple (BS2) avec les sorties correspondantes (out2)

La courbe obtenue après implémentation sous MATLAB montre une amélioration des performances puisque l'erreur quadratique est de 0.130864 (SSE=0.130864) pour un objectif de 0.1 (Goal=0.1) au bout de 2000 itérations alors qu'en phase d'apprentissage 0.09 au bout de 958 itérations.

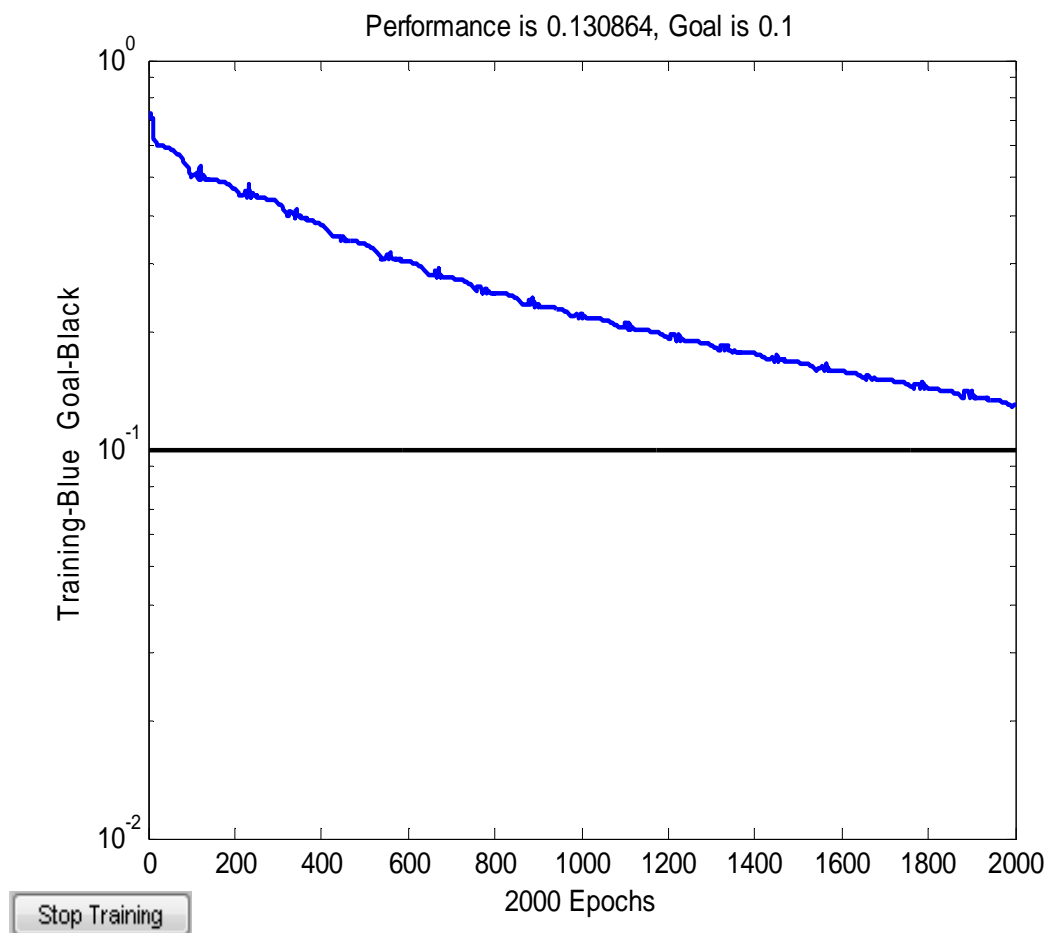


Figure IV-3 : Performance du réseau 2 de la 2^{ème} base d'exemple

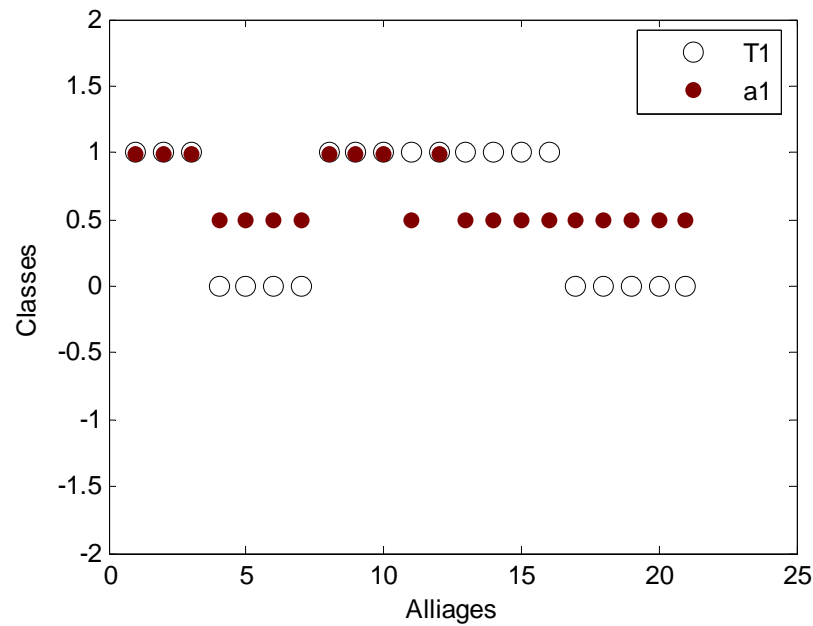


Figure IV-4 Représentation du nombre d'alliages classés par le net2

La sortie (a1) du net2 et la sortie prévue T1 sont représentées sur la figure (IV-4). Les alliages dont la teneur en cuivre est supérieure à 2% et précisément une teneur de 2.6%, leur classification par le net et celle prédite sont en accord. En d'autres termes ces alliages appartiennent à la classe (1) donc ils sont durs et fragiles avec un G qui varie entre (25.858GPa et 27.273GPa) prédits par le net.

Les alliages présentant une teneur en cuivre inférieure à 2% (dure et ductile) leur classification par le net2 ne s'accorde pas avec la classification supervisée, avec un module de cisaillement de (G=27.162GPa) et un module de compression (B=70.124GPa) Dans cette étape donc seules les alliages présentant une teneur en cuivre supérieure à 2% leurs la classification supervisée et prédite sont en accord.

IV-5-3 Test de la classification du net2

Dans cette étape on va tester la classification du réseau de neurones artificiels (net2), pour cela on présente au réseau en entrée la première base de données (BS1) indiquée sur le tableau (IV-1), auquel on ajoute un bruit de 20%, c'est-à-dire ces données seront multipliées par 0.2

$$P = P1 + \text{randn}(R, Q)*0.2$$

La figure (IV-5) obtenue montre à quel point les performances sont encore bonnes, l'erreur quadratique est de (SSE=0.0995503) au bout de 902 itérations.

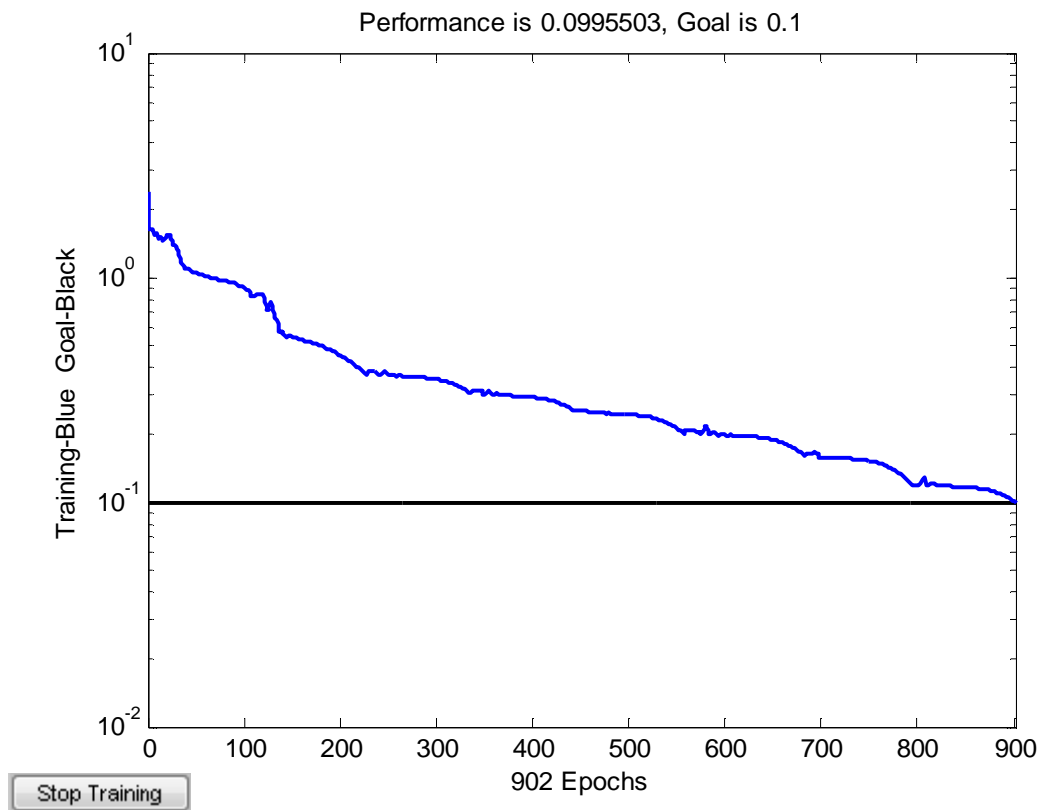


Figure IV-5 : Performance du réseau 2 bruité

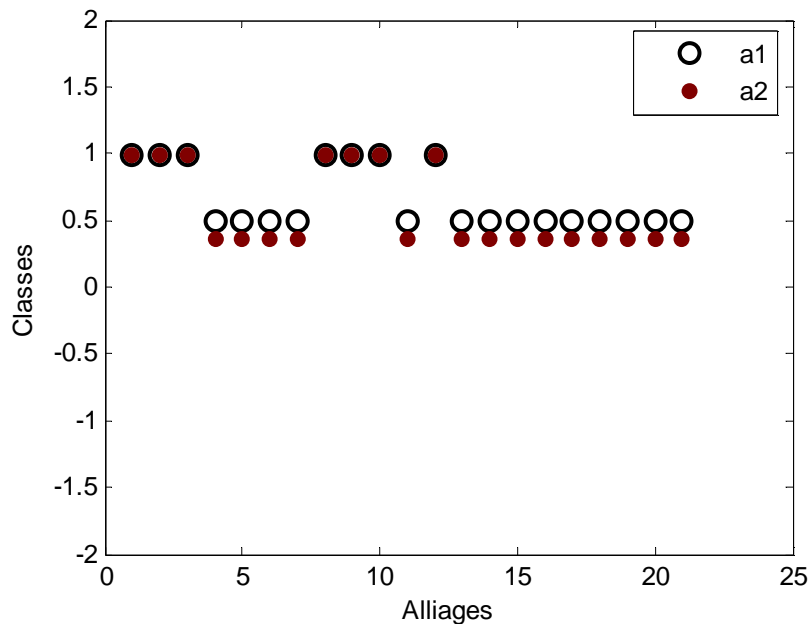


Figure IV-6 Représentation de la classe prédite et désirée

Dans cette étape de classification on montre à quel point le net classe les données même en présence de bruit.

Il faut penser que le but de cette étape est d'arriver à concevoir un réseau qui serait tolérant aux bruits au risque où les données acquises en réalité sont difficiles à avoir avec une précision exacte.

Cette phase de conception est très importante car le réseau doit être préparé de sorte qu'il supporte ce genre de perturbations d'un côté pour minimiser les erreurs sur la classification et d'un autre avoir des prédictions pour de nouveaux alliages.

On simule encore une classification des données de la base (BS1) avec le net3 version entraînée avec bruit, la sortie étant a2. Les sorties du net sont les classes binaires. La différence ($e1-e2$) montre l'erreur de classification des mêmes données entre net1 et net3.

L'erreur totale de classification qui est l'erreur quadratique minimale (MSE) est donnée à la fin ; plus cette erreur (s) est faible plus le net est performant en terme de classification. Les sorties désirées ($out1$) et les sorties obtenues par le réseau de neurones sont illustrées sur le tableau IV-4.

a1	Out1	a1	Out1
1	1	0,76923077	1
1	1	0,76923077	1
1	1	0,76923077	1
0,03846154	0	0,76923077	1
0,03846154	0	0,76923077	1
0,03846154	0	0,76923077	0
0,03846154	0	0,76923077	0
1	1	0,76923077	0
1	1	0,76923077	0
1	1	0,76923077	0
0,76923077	1		

Tableau IV-4 : Représentation des sorties obtenues par RNN (a1) et désirées (out1)

On remarque que parmi les 21 échantillons (alliages) seulement 5, dont la classification supervisée et prédite ne sont pas en accord (le tableau IV-5) .Dans ce cas on peut dire que les paramètres données en entrée, les différentes concentrations des éléments d'addition et les propriétés mécaniques doivent être ajustés. Pour avoir un matériau avec des propriétés spécifiques telle que la dureté et la ductilité en même temps il faut prévoir un alliage dont la teneur en cuivre soit inférieure à 1.90% .

7475 T61	88,5	0,25	1,90	0,12	2,60	0,06	0,10	0,06	6,20	148	2,81	11	68,9216	27,00	2,5527
7475 T65	88,5	0,25	1,90	0,12	2,60	0,06	0,10	0,06	6,20	150	2,81	13	70,2941	27,00	2,6035
7475 T761	88,5	0,25	1,90	0,12	2,60	0,06	0,10	0,06	6,20	140	2,81	12	68,9216	27,00	2,5527
7475 T7351	88,5	0,25	1,90	0,12	2,60	0,06	0,10	0,06	6,20	135	2,81	13	70,2941	27,00	2,6035
7475 T7651	88,5	0,25	1,90	0,12	2,60	0,06	0,10	0,06	6,20	140	2,81	12	70,2941	27,00	2,6035

Tableau IV-5 Les alliages dont la classe non accord avec le net.

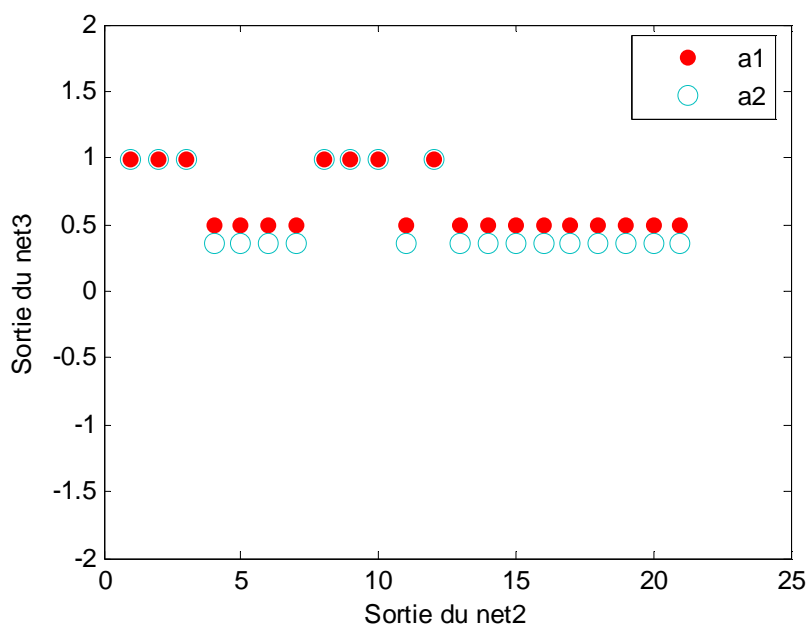


Figure IV -7 Nombre d'alliage classés par le net3

La différence de classification des données (BS1) entre le net2 (sans bruit) et le net3 (avec bruit), sont représentées sur la figure (IV-4). On constate que même si on apporte un bruit sur le réseau d neurones, les sorties se rapprochent.

Cela pour dire que le réseau a été bien entraîné dans la phase apprentissage et qu'il a été bien conçu. Dans ce cas on peu généraliser cette classification à d'autres alliages en instance.

IV-5-4 Prédiction des propriétés élastiques de l'alliage Al-Zn

L'utilisation des réseaux de neurones dans la classification des alliages d'aluminium Al-Zn nous a permis de déterminer dans cette partie du travail certaines propriétés comme le module de compression B, le module de cisaillement G et le ratio B/G donnés par le net en fonction des entrées présentées au net. D'après la représentation graphique du module de compression prédit en fonction de l'expérimental, figure (IV-5) nous pouvons distinguer que les deux alliages Al-Zn(7175) et le Al-Zn(7001) présentant respectivement une teneur en cuivre de 2% et 2.6% appartiennent à la classe «dure et fragile» s'ils présentent un module de compression de 71.004GPa et 69.694GPa.

Dans le cas du module de cisaillement comme indiqué sur la figure (IV-6), L'alliage Al-Zn (7005) dont la teneur en cuivre est de 0.1% et un module de cisaillement appartient à la classe « dure et ductile » s'il présente un $G=27.097\text{GPa}$ au lieu de 26.90 GPa.

La figure (IV-7) représente le ratio B/G prédit par le net en fonction de l'expérimental. L'alliage Al-Zn(7005) avec une teneur en cuivre de 0.1% est ductile si son ratio est de 2.9765 qui est la valeur maximale de ces alliages.

B (GPa) prédit	B (GPa) expérimental	B (GPa) prédit	B (GPa) expérimental
69,608	69,563	70,294	70,474
69,608	69,285	70,588	70,278
69,608	69,694	70,588	70,17
70,588	70,934	70,588	70,636
70,588	70,492	70,588	71,004
68,235	67,998	68,922	69,073
68,235	67,924	70,294	70,064
70,294	70,559	68,922	68,761
70,294	70,4	70,294	70,517
70,294	70,358	70,294	70,206
70,294	69,968		

Tableau IV-6 Les modules de compressibilité prédit et calculé

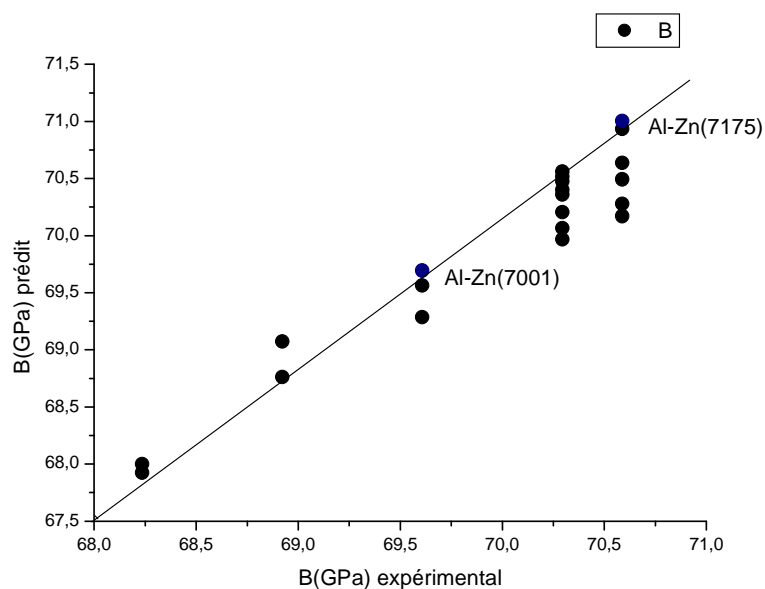


Figure VI-8 Représentation du B prédit et l'expérimenta

G (GPa) prédit	G (GPa) expérimental	G (GPa) prédit	G (GPa) expérimental
27,114	27	26,828	26,9
26,754	27	27,156	27
26,853	27	26,394	27
27,042	26,9	27,59	27
27,097	26,9	26,853	27
25,845	26	26,932	27
25,745	26	26,671	27
25,959	26,9	27,177	27
26,875	26,9	26,998	27
26,789	26,9	27,112	27
26,87	26,9		

Tableau IV-7 Les modules de cisaillement prédit et calculés

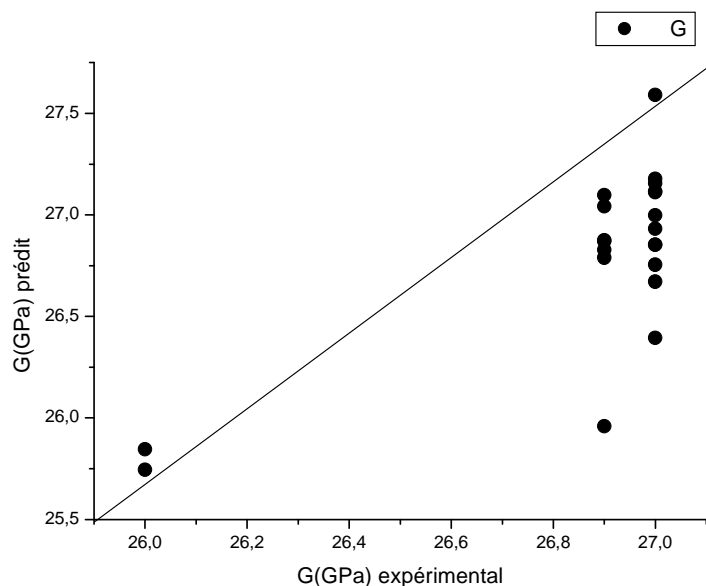


Figure IV-9 Représentation de G prédit en fonction de la valeur d'entrée

B/G prédit	B/G expérimental	B/G prédit	B/G expérimental
2,518	2,5781	2,5875	2,6132
2,6195	2,5781	2,7225	2,6144
2,6911	2,5781	2,8856	2,6144
2,4745	2,6241	2,2585	2,6144
2,9765	2,6241	2,7604	2,6144
2,6786	2,6244	2,3413	2,5527
2,5393	2,6244	2,5468	2,6035
2,4657	2,6132	2,4073	2,5527
2,6728	2,6132	2,6992	2,6035
2,6763	2,6132	2,619	2,6035
2,5336	2,6132		

Tableau IV-8 Le ratio prédit et calculés

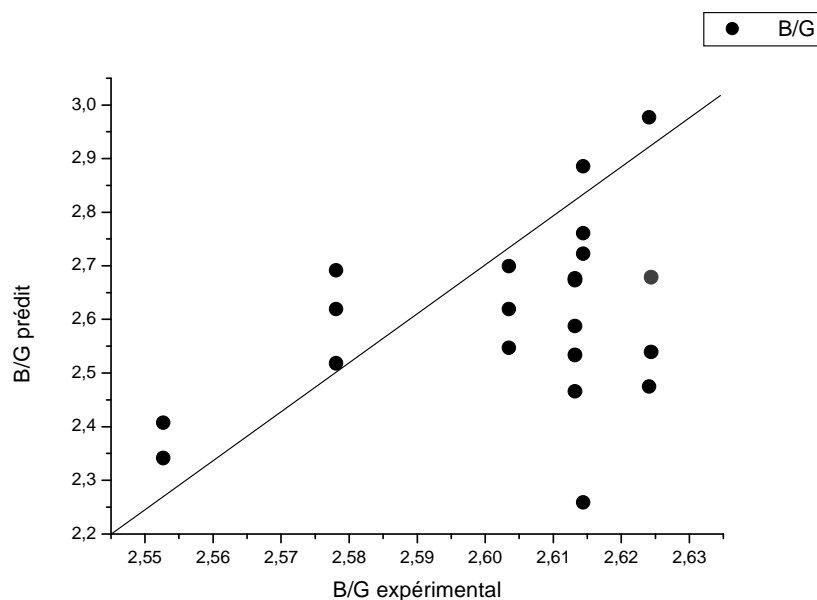


Figure IV-10 Représentation du ratio (B/G) prédit et expérimental

IV-5-5 Généralisation

Enfin pour une généralisation on peut utiliser le net comme classifieur, il suffit de lui faire soumettre n'importe quel vecteur de dimension 14 en d'autres termes lui présenter en entrées 14 paramètres d'un alliage quelconque, le réseau de neurones artificiels est capable de nous donner en sortie un bit 0 si classe "dure et ductile", ou bit 1 si classe "dure et fragile". Pour cette analyse nous avons présenté un alliage avec une teneur de 2.6% en cuivre pour faire la simulation. Les sorties obtenues par le net2 c'est-à-dire $a_1=0.99189$ et la sortie du net3 est $a_2=0.9995$, avec une erreur minimale ce qui veut dire que cet alliage est dur et fragile. Les concentrations des éléments d'additions et les propriétés de cet alliage sont donnés sur le tableau IV-9.

		Données en entrée	Sortie du net
Eléments d'addition	Al	0,87	1,0378
	Cr	0,35	0,20906
	Fe	0,4	0,51004
	Mg	3,4	3,7359
	Mn	0,2	0,20938
	Si	0,35	0,42803
	Ti	0,2	0,28756
	Zn	8	7,9202
Propriétés	dureté	60	59,893
	densité	2,84	2,5685
	A%	14	13,983
	B	69,608	69,553
	G	27	27,093
	B/G	2,5781	2,4646

Tableau IV-9 Propriétés obtenues par le net de l'alliage Al-Zn 7001

Nous avons présenté un autre alliage d'aluminium mais avec une teneur de 0.1% en cuivre, donc la classe sera « dure et ductile ».les sortie du net2 et net3 sont respectivement 0.02639 et 0.01634, ce qui est en accord avec la classe voulue et les concentrations des éléments d'addition de cet alliage et ses propriétés après ajustement par le réseau de neurones artificiels sont indiqués sur le tableau IV-10.

		Données en entrée	Sortie du net
Eléments d'addition	Al	91,113	91,035
	Cr	0,3912	0,00089989
	Fe	0,44557	0,54957
	Mg	1,7874	2,0663
	Mn	0,73223	0,83425
	Si	0,42405	0,31238
	Ti	0,023922	0,068201
	Zn	4,9744	4,7269
Propriétés mécaniques	dureté	104,84	105,38
	densité	2,6981	2,6666
	A%	15,021	14,991
	B GPa	70,299	70,326
	G GPa	27,053	27,158
	B/G	2,7584	2,7404

Tableau IV-10 Propriétés de l'alliage (%Cu=0.1) par le RNN

IV-5 Conclusion

L'utilisation des réseaux de neurones pour la classification des alliages d'aluminium suivant leurs deux propriétés mécaniques ductile et fragile et l'outil informatique avec, montrent à quel point cette méthode apporte de l'aide à la décision pour déterminer ce qui est nécessaire et utile dans le but de la conception de nouveaux matériaux.

Nous avons utilisé un alliage d'aluminium de la série 7000 dont l'élément d'addition principale est le zinc pour une classification bien définie dite supervisée afin de pouvoir étudier les propriétés mécaniques de ces matériaux.

Le résultat était assez performant en termes d'erreur sur la classification et aussi prédiction de nouvelles données pour la conception de nouveaux matériaux. Cette technique du data mining s'avère très encourageante dans l'étude des matériaux.

Le perceptron multicouche comme modèle de réseau de neurones artificiels à apprentissage supervisée permet d'avancer pour envisager de classer de très grandes bases de données et de les traiter dans un temps record.

Nous avons adopté une programmation sous MATLAB des réseaux de neurones et le programme est le suivant :

IV-6 Programme

%physique: NN for classifying physical materials.

```

clf;
figure(gcf)
echo on

% NEWFF - Initializes feed-forward networks.
% TRAINgdx - Trains a feed-forward network with faster backpropagation.
% SIM - Simulates feed-forward networks.
pause % Strike any key to continue...

% load data samples of materials
ndataIn = xlsread ('C:\files\BASES.xlsx','BS1');
ndataOut = xlsread ('C:\files\BASES.xlsx','out1');
P1=ndataIn'; % each column is an input node for the NNet, nber of rows is the nber
of sample data.
T1=ndataOut'; % in this case 1 output node (binary 0 or 1)
[R,Q] = size(P1);
[S2,Q] = size(T1);
pause % Strike any key to define the network...

% DEFINING THE NETWORK:
% nber of input nodes= nber of columns (nber of parameters that define a
material);
% S1 nber of nodes in the hidden layer; and S2 is the nber of nodes in
% the output layer (output real values for 2 classes:
% Classe (1): "dures et ductiles" avec la teneur en cuivre inférieure à 2%
% Classe (2) :bit 1: "dures et fragiles" avec la teneur en cuivre supérieure ou égale à
2%
% =====

S1 = 14;
net = newff(minmax(P1),[S1 S2],{'logsig' 'logsig' },'traingdx');
% feed forward NN with R input nodes; S1 hidden nodes and S2 output nodes;
% all nodes are of sigmoid transfert functions; training with accelerated back
propagation algorithm.
pause % Strike any key to train the network...

```

```

% TRAINING THE NETWORK WITHOUT NOISE
% =====
% NN training parameters

net.performFcn = 'sse';           % Sum-Squared Error performance function
net.trainParam.goal = 0.1;       % Sum-squared error goal.
net.trainParam.show = 20;       % Frequency of progress displays (in epochs).
net.trainParam.epochs = 4000;   % Maximum number of epochs to train.
net.trainParam.mc = 0.95;       % Momentum constant.
% Training begins using the data samples input/output P1-T1...please wait...
[net,tr] = train(net,P1,T1);
% this uses training data samples; the figure obtained shows a good
% convergence of error to a min value (target non reached)

% Now, a copy of this net is made (with same structure and as being already
% trained):
net2=net;
% this time this net will be trained with test data (2nd data sets)
ndataIntest = xlsread('C:\files\BASES.xlsx','BS2');
ndataOuttest = xlsread('C:\files\BASES.xlsx','out2');
P2=ndataIntest' ;
T2=ndataOuttest' ;
pause; % strike key to go on...
% training....
[net2,tr] = train(net2,P2,T2);
% the figure obtained must show an improvement in training performance

% here we try simulation of classification of some alloys (2nd data base
% for example)
x1=P1;
pause % Strike any key to see the output for these materials (classes to which they
should belong) :
a1=sim(net2,x1);
% Now in this ste; we try to take into consideration noise that may affect any
measures
% of the alloys parameters (during acquisition process)
pause; % Strike any key to train the network with noise

```

```

% TRAINING THE NETWORK WITH NOISE
% =====

net3 = net;
net3.trainParam.goal = 0.1;           % Mean-squared error goal.
net3.trainParam.epochs = 3000;       % Maximum number of epochs to train.
% The network will be trained with noisy data added to the 2nd data base
[R,Q] = size(P1);
P = P1 + randn(R,Q)*0.2;             %random noise (20%) added to the sample data to
simulate noisy data
pause % Strike any key to begin training...

net3.performFcn = 'sse';              % Sum-Squared Error performance function
net3.trainParam.goal = 0.1;          % Sum-squared error goal.
net3.trainParam.show = 20;           % Frequency of progress displays (in epochs).
net3.trainParam.epochs = 3000;       % Maximum number of epochs to train.
net3.trainParam.mc = 0.95;           % Momentum constant.
% Training begins...please wait...
[net3,tr] = train(net3,P,T1);
pause % Strike any key to see the output for these materials (vector x1):
a2=sim(net3,x1);

% error (performance of classification even in presence of noise)
e=a1-a2;
s=mse(e) % sum squared error; the less this mean squared error (mse) the best are
the performance of the NNet as a classifier.

%h we may try any vector x (alliage); with sim(net3,x) that gi which gives the
%class to which it belongs: 0 or 1 (see the coding above).
echo off
disp('End of phys1')

```

Conclusion

Conclusion générale

Ce mémoire s'inscrit dans la classification par l'approche d'une technique de data mining, les réseaux de neurones artificiels. Les techniques et méthodes du data mining sont nombreuses et offrent de bonnes perspectives pour l'extraction de connaissances à partir de grandes bases de données.

Une contribution dans l'application de l'une des techniques du data mining vient de la classification des matériaux suivant leurs propriétés élastiques en adoptant un nouveau algorithme d'apprentissage automatique qui est le perceptron multicouche comme modèle de réseau de neurones artificiels.

Pour la réussite d'un projet de data mining il faut des bases de données, dans un premier temps nous avons construit une base de données d'un alliage d'aluminium de la série 7000 dont l'élément d'addition est le zinc en donnant les concentrations des éléments d'addition et quelques propriétés élastiques.

Dans une deuxième partie, nous présentons un modèle de Réseaux de Neurones pour la classification de ces alliages suivant deux classes appropriées à cette série d'alliage qui sont « dure ductile » et « dure fragile » en tenant compte des teneurs de l'un des éléments d'addition, le cuivre.

La base de donnée est divisée en deux, une première pour l'apprentissage du réseau et comme entrées 21 alliages avec leurs concentrations et leurs propriétés et une sortie supervisée codée par 0 pour dure ductile et 1 pour dure fragile.

Le modèle de réseaux de neurones, bien entraîné a donné de bonnes performances en termes d'erreur sur la classification.

Nous avons ensuite présenté une deuxième base de donnée ayant la même structure que la base d'apprentissage avec une variation dans les concentrations en aluminium et la aussi les sorties du net sont en accord avec ceux voulu et avec un minimum d'erreur.

Pour le test de la phase d'apprentissage il faut simuler un bruit sur la bases d'apprentissage en d'autres termes perturber les données avec un certains pourcentage, dans notre cas la perturbation est de 20% .les sorties du net en terme de classifieur sont performantes a juger par l'erreur quadratique qui est minimale même en présence de bruit.

Le résultat de calcul de la différence entre l'erreur de classification avec le (net2) sans bruit et le (net3) avec bruit a été assez minimale et donc bonnes performances De classification.

La dernière étape de ce programme ou plus précisément a la fin de l'algorithme il faut généraliser en présentant au réseau un alliage avec le même nombre de paramètres que ceux donner en apprentissage.

Nous avons donné deux alliages l'un avec une teneur de 2.6% en cuivre et un deuxième avec 0.1%.

Les sorties obtenues par le réseau sont en accord avec les classes spécifiées et avec un minimum d'erreur, mais aussi après ajustement de ses poids le réseau de neurones a fourni les concentrations des différents éléments constituant l'alliage dans le but de préciser la classe dont il doit appartenir et par la suite le domaine d'utilisation de ce matériau , donc la conception d'autres alliages

L'adoption de cette technique de classification a permis de montrer aussi a quel point l'outil informatique est nécessaire pour la réussite de telles études et l'informatique des matériaux s'avère prometteuse pour la prédiction des matériaux

Le calcul par ces nouvelles méthodes et approches de data mining produit des résultats proches de l'expérimentale, chose assez difficile a réaliser en terme de cout et temps de calcul.

On peut conclure que l'utilisation des outils de data mining nombreuses quelles sont et parmi eux les réseaux de neurones a beaucoup contribuer à la génération de

modèle comportementale comme le comportement mécanique des alliages d'aluminium de la série 7000 qui présentent une dureté importante.

Les Réseaux de Neurones constituent des outils d'investigation puissants pour différents domaines d'application en Sciences des Matériaux, ces méthodes connaissent actuellement un essor considérable dans le domaine de l'optimisation et la prévision de comportement grâce à leur universalité, fiabilité et robustesse.

Références

- [1] : Marie Baltzer Juin 2002 DESS 220 Informatique de gestion Data mining : Etude des 100 premiers groupes mondiaux.
- [2] : Jiawei Han, Micheline Kamber. Data Mining . Concepts and Techniques
- [3] : Georges El Helou et Charbel Abou khalil, Datamining, techniques d'extraction de connaissances, projet soutenu le 16/02/2004
- [4] : C.K.R. Suh, B.M. Vogel, B. Narasimhan, S.K. Mallapragada, Combinatorial Materials Science, Eds. S.K. Mallapragada, B. Narasimhan, and M.D. Porter 2007, Hoboken, NJ, John Wiley-Inter science.
- [5] : Ruggero G. Projet "K-Means". 16 novembre 2006
- [6] : Ricco Rakotomalala, Arbres de Décision, Laboratoire ERIC
Université Lumière Lyon 25, av. Mendés France, e-mail : rakotoma@univ-lyon2.fr
- [7] : Stéphane Tufféry . DATA MINING & STATISTIQUE DECISIONNELLE
<http://data.mining.free.fr>
- [8] : G. Saporta, Conservatoire National des Arts et Métiers ; <http://Cedric.Cnam.fr/Saporta>
- [9] : Philippe Besse, Alain Baccini. Référence AF620 | Date de publication : 10 avril 2011
- [10] : Sandro.Saitta@epfl.ch, 'Data Mining des données à la connaissance' EPFL/IMAC (Laboratoire d'informatique et de mécanique appliquée à la construction).
- [11] : Rabah Taouche 'Prévision du comportement mécanique d'alliages biphasés par algorithmes génétiques et réseaux de neurones'. application aux systèmes WC-CO: 19/01/2010
- [12] : Manuel Clergue. Réseau de neurones artificiel.
- [13] : Marc Parizeau Le perceptron multicouche et son algorithme de rétropropagation des erreurs .Département de génie électrique et de génie informatique Université Laval. 10 septembre 2004
- [14] : ROS. Le neurone biologique
- [15] : G. Dreyfus 'Les réseaux de neurones', Ecole Supérieure de Physique et de Chimie Industrielles de la Ville de Paris (ESPCI), Laboratoire d'Électronique
- [16] : Yoshua Bengio 'Algorithmes d'apprentissage' www.iro.umontreal.ca/~pift6266

[17] : Bessai F.Z., Krelifaoui Mohamed, Guergab Mohamed centre de Recherche sur l'Information Scientifique et Technique (CERIST) Division Base de Données et Système multimédia zbessai@wissal.dz

[18] : Vincent Lemaire .Les réseaux de neurones les réseaux de neurones

[19] :Ph. Leray <http://servasi.insa-rouen.fr/~pleray/ftp/> Ph. Leray

[20] : Claude Touzet ‘ Les réseaux de neurones artificiels introduction au connexionnisme’.

[21] : D. Urbani, P. Roussel-Ragot, L. Personnaz, G. Dreyfus, ‘The selection of neural models of nonlinear dynamical systems by statistical tests’, Neural Networks for Signal Processing I, pp. 229-237, IEEE Press, 1994.

[22] : Fabien Moutarde .Introduction aux réseaux de neurones. Centre de Robotique (CAOR) Mines ParisTech (ENSMP)

[23] : SAHA Ouafi1* .Boutaleb ali*.Applications des réseaux de neurones aux matériaux dans le génie civil état de l’art.

Ecole Normale Supérieure de l’Enseignement Technique d’Oran

[24] : A. Bahrami, S. H. MousaviAnijdan, A. Ekrami, ‘Prediction of mechanical properties of DP steels using neural network model’, Journal of Alloys and Compounds, 392, pp. 177-182, 2005

[25] : J. Holland, ‘Adaptation in Natural and Artificial Systems’, Ann Arbor, University of Michigan Press, 1975.

[26] : Fabrice SORIN,Lionel Broussard et Pierre Roblin, Régulation d'un processus industriel par réseaux de neurones,Document délivré le21/05/2012 Pour le compte7200092269 - CERIST // 193.194.76.5

[25] : Jonathan Idrac, COMPORTEMENT ELECTROCHIMIQUE D’ALLIAGES MODELES BINAIRES ALUMINIUM-CUIVRE, THESEDE DocteurDE L’INSTITUT NATIONAL POLYTECHNIQUE DE TOULOUSE.

[26] : J. Manec, « Thermodynamique structural des alliages». Presse universitaire deFrance (1972).

[27] : A. RASKIN,Rappels atomistiques, structure des métaux, des alliages et des céramiques, Université Médicale Virtuelle Francophone Date de création du document 2009-2010

[28] : F.D.Geuser. Thèse de doctorat, Université de ROUEN (2005).

- [29] :A.Raskin, Rappels atomistiques, structures des métaux, des alliages et des céramiques. Université médicale virtuelle francophone, 2009-2010
- [30] : Azizi Assia, Etude de l'évolution microstructurale des alliages ou se forment les zones de Guinier-Preston, Mémoire présenté pour obtenir le diplôme de Doctorat en Sciences.
- [31] : I. Meunier, Thèse de Doctorat, Université de Marseille, (2001).
- [32] : J. Hertz, «Diagrammes d'équilibre Alliages binaires», Technique de l'ingénieur», M70, (1999).
- [33] :François Hamon, *Modélisation du comportement mécanique en fissuration d'alliages aéronautiques*, doctorat Soutenu le 10 mai 2010 .
- [34] : A. Pasturel, Thèse de Doctorat, Université de Grenoble (1983).
- [35] : Y. Berthaud «Matériaux et propriétés», Université Pierre et Marie Curie, (2004).
- [36] :C. Kittel, «Physique de l'état Solide», Edition, Dunod, (1983).
- [37] : Jonathan Idrac, COMPORTEMENT ELECTROCHIMIQUE D'ALLIAGES MODELES BINAIRES ALUMINIUM-CUIVRE,Thèsede docteurde l'institut national polytechnique de Toulouse.
- [38] : J. Manec, « Thermodynamique structural des alliages». Presse universitaire deFrance (1972).
- [39] : I. Meunier, Thèse de Doctorat, Université de Marseille, (2001).
- [40] : J. Hertz, «Diagrammes d'équilibre Alliages binaires», Technique de l'ingénieur», M70, (1999).
- [41] : A. Pasturel, Thèse de Doctorat, Université de Grenoble (1983).
- [42] : Y. Berthaud «Matériaux et propriétés», Université Pierre et Marie Curie, (2004).
- [43] :Benyelloul Kamel , *Application du Dataminingpour la prédiction et l'optimisationd'un nouvel alliage d'Aluminium de la série7xxx*
- [44] :Roger DEVELAYPropriétés de l'aluminium et des alliages d'aluminium corroyés ,Référence M440|Date de publication : 10 juil. 1992 |

Résumé

Ces dernières années ont vu le développement des techniques de fouille de données dans de nombreux domaines d'applications dans le but d'analyser des données volumineuses et complexes. La génération de modèle de comportements de processus physique à partir de données font l'objet de cette analyse en se servant des réseaux de neurones comme technique.

Ce travail consiste à la classification supervisé de données physiques d'un matériaux, que nous nous proposons d'étudier à travers une nouvelle approche basée sur le data mining. Il s'agit de mettre en place une nouvelle approche de classification automatique basée sur une technique de la théorie des réseaux de neurones. Le matériau étudié est un alliage d'aluminium de la série 7000 classé suivant deux propriétés ductile ou fragile en tenant compte des éléments d'addition. Le module de compressibilité (B), le module de cisaillement (G), et le rapport B/G sont utilisés pour étudier le comportement de ce matériau.

Nous avons également développé l'apprentissage incrémental associé à cette approche, ce qui permet à de nouvelles données d'être automatiquement intégrées dans la partition initialement générée sans avoir à relancer la classification globale. Le modèle obtenu assure une bonne interprétabilité des données de l'alliage d'aluminium (Al-Zn) construites et permet d'autre part de prédire la classe d'un nouvel alliage et donc ses propriétés élastiques.

Mots clés : Datamining ; Réseaux de neurones ; Classification ;
Perceptron multicouches ; Alliage Al-Zn.

Abstract

Recent years have seen the development of data mining techniques in many application areas in order to analyze large and complex data. Generation model behavior from physical processes is the subject of this analysis using neural networks as a technique.

This work involves the supervised classification of physical data of a material; we propose to study through a new approach based on data mining. This is to introduce a new approach to automatic classification based on a technical theory of neural networks. The studied material is an aluminum alloy of the 7000 series classified according to two ductile or brittle properties taking into account the alloying elements. Bulk modulus (B), the shear modulus (G), and the ratio (B/G) are used to study the behavior of this material are used to study the behavior of this material.

We have also developed incremental learning associated with this approach, allowing new data to be automatically included in the initial partition generated without having to restart the overall classification. The resulting model provides a good interpretability of data from the aluminum alloy (Al-Zn) and constructed on the other hand allows to predict the class of a new alloy and therefore its elastic properties.

Keywords: Data mining; Neural networks; Classification;
Multilayer Perceptron; Alloy Al-Zn

ملخص

شهدت السنوات الأخيرة تطور تقنيات ما يسمى (Datamining) في مجالات عدة قصد تطبيقها في تحليل المعطيات الكثيرة والمعقدة. من أجل توليد نموذج مثالي للظواهر الفيزيائية عن طريق المعطيات تم تطبيق تقنية جديدة وهي (réseaux de neurones artificiels).

يتخصص هذا العمل في تصنيف مرغوب و موجه لمعطيات فيزيائية للمادة قصد التحليل للمعلومة المتحصل عليها و التصنيف الالي في هذا الاطار. المعلومات المستعملة هي المواد الاضافية وبعض الخواص الفيزيائية كالصلابة والمرونة.

المادة المقصود تحليل خواصها هي سبائك الألمنيوم (Al-Zn) من الصنف 7000. تم تصنيفها حسب حلتين المرونة و الصلابة. في الأخير وبعد استعمال تقنية (Réseau de neurones) تحصلنا علي نتائج تساعد في تصنيف اي مادة معطاة و التنبؤ بخواصها الفيزيائية.