

N°d'ordre : 06/2021 – D/PH

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE

Ministère d'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITÉ DES SCIENCES ET DE LA TECHNOLOGIE

HOUARI BOUMEDIENNE U.S.T.H.B

Faculté De Physique



Thèse de Doctorat en Sciences

En vue de l'obtention du Grade de

Docteur en **Physique** Spécialité : **Physique Théorique**

par :

Sanaa Zaoui épouse Melab

Thème :

Application des techniques neuronales pour la prédiction des propriétés thermoélectroniques des éléments.

Soutenue publiquement, le 07/04/2021, devant le jury de thèse :

M. Abdelhafid KELLOU	Professeur à l' USTHB	Président
Mme. Fawzia MEKIDECHE-CHAFA	Professeur à l' USTHB	Directrice de Thèse
M. Nasser BENREKAA	Professeur à l' USTHB	Examinateur
M. Boualem BOURAHLA	Professeur à l' UMMTO	Examinateur
M. Salah BLIZAK	Maître de conférence/A à l' UMBB	Examinateur
M. Mustapha SADOUKI	Maître de conférence/A à l' UDBKM	Examinateur
M. Adel BELAYADI	Maître de conférence/B à l' ESSAIA	Invité

REMERCIEMENT

Le travail de recherche présenté dans ce manuscrit a été accompli au sein du Laboratoire de Physique Théorique de la Faculté de Physique de l'Université des Sciences et de la Technologie Houari Boumediene d'Alger.

Tout d'abord, je tiens à exprimer mes remerciements à ma directrice de thèse, Madame F. MEKIDECHE-CHAFA, Professeur à l'USTHB, je tiens à lui exprimer ma gratitude pour avoir dirigé ce travail, pour son encadrement sans faille, ses conseils judicieux, sa constante disponibilité, son souci de clarté et toutes les remarques avisées qu'elle a pu me transmettre.

Toute ma reconnaissance et profonds remerciements vont à Monsieur A. Belayadi, pour son soutien, son enthousiasme, sa vision et son aide précieuse qui m'a fait avancer et m'a permis d'accomplir ce travail. Une fois de plus un grand merci.

Je tiens à remercier vivement Monsieur A. Kellou, Professeur à l'USTHB, pour avoir accepté de présider mon jury et d'évaluer mon travail.

Je souhaite, de plus, remercier les membres du jury, les Docteurs S. Blizak Maître de conférences à l'Université Mohamed Bouguera à Boumerdes et M. Sadouki, maître de conférences à l'Université Djilali Bounaama à Khemis Miliana qui m'ont fait l'honneur de contribuer à l'examen de mon travail.

Mes plus vifs remerciements vont à Monsieur B. Bourahla, Professeur au Laboratoire de physique et chimie quantique de l'Université Mouloud Mammeri de Tizi-Ouzou qui a accepté de participer à mon jury de soutenance.

Je tiens, par ailleurs, à exprimer mes plus sincères remerciements à Monsieur N. Benrekaa, Professeur à l'USTHB qui nous a honorés par sa présence en acceptant de faire partie du jury de thèse.

Enfin, je ne saurais oublier de remercier infiniment mes parents pour leur amour, leur dévouement et leur soutien. Qu'ils trouvent ici la récompense à tous leurs sacrifices et toute ma reconnaissance. Un énorme merci également à mon mari, mes enfants, mon frère et à mes soeurs pour avoir toujours été présents à mes côtés et pour cette fratrie unie que nous formons.

Table des matières

Introduction Générale	11
1 Chapitre 1: Généralités sur les réseaux de neurones	17
1.1 Introduction	18
1.2 Inspiration biologique	19
1.3 Historique du neurone artificiel (ou neurone formel)	21
1.4 Le neurone formel de McCulloch et Pitts	22
1.5 Les fonctions d'activation	23
1.5.1 Fonction seuil	23
1.5.2 Fonction linéaire	24
1.5.3 Fonction linéaire par morceaux (Piecewise)	25
1.5.4 Fonction sigmoïde	26
1.5.5 Fonction radiale	27
1.6 Interconnexions neuronales	28
1.7 Apprentissage dans les réseaux de neurones	30
1.7.1 Apprentissage supervisé	30
1.7.2 Apprentissage non supervisé	31
1.7.3 Les règles d'apprentissage	32
1.8 Modèles à base de réseaux de neurones	34
1.8.1 Les réseaux à propagation avant	34
1.8.2 Les réseaux à interconnexion complète	41

1.9	Conclusion	45
2	Chapitre 2 : Méthodologie des réseaux multicouches et algorithmes d'optimisation	46
2.1	Introduction	47
2.2	Méthodologie des réseaux multicouches et algorithmes d'optimisation	47
2.3	Algorithme d'apprentissage	48
2.3.1	Méthodes du premier ordre	49
2.3.2	Méthodes du second ordre	55
2.4	Intérêt des réseaux de neurones formels	60
2.5	Propriété fondamentale des réseaux de neurones artificiels	61
2.6	Les différents domaines d'application des réseaux de neurones	63
2.7	Conclusion	65
3	Chapitre 3 : La thermoélectricité	66
3.1	Introduction	67
3.2	Présentation	68
3.3	Effets thermoélectriques	71
3.3.1	L'effet Seebeck	71
3.3.2	L'effet Peltier	71
3.3.3	L'effet Thomson	72
3.3.4	Les relations de Kelvin	73
3.4	Conversion d'énergie	74
3.5	Rendement	76
3.6	Technique de mesure des propriétés thermoélectriques	78
3.6.1	La réponse thermique d'un dispositif thermoélectrique	78
3.6.2	Calcul de la fonction de l'impédance dans le domaine des fréquences	79
3.6.3	Influence du contact métallique	81
3.7	Conclusion	88

4 Chapitre 4 : Les réseaux de neurones pour prédire les propriétés thermoélectriques des éléments	89
4.1 Introduction	90
4.2 Modélisation des propriétés thermoélectriques des éléments par l'approche des réseaux neuronaux	92
4.2.1 Configuration de la structure du réseau	93
4.2.2 Mise en place de la phase d'apprentissage du réseau	95
4.2.3 Validation du modèle par représentation de Nyquist	98
4.3 Validation du modèle via la réponse du réseau	108
4.4 Conclusion	111
Conclusion Générale	112
Bibliographie	116

Table des figures

1.1	Le neurone biologique [2]	20
1.2	Schématisation algorithmique d'un neurone artificiel [9].	22
1.3	Représentation de la fonction seuil.	24
1.4	Représentation de la fonction linéaire.	25
1.5	Représentation de la fonction linéaire par morceaux ($\theta = 40$).	26
1.6	Représentation de la fonction sigmoïde logistique (logsig).	26
1.7	Représentation de la fonction sigmoïde tangente hyperbolique (tangsig).	27
1.8	Représentation de la fonction radiale (gaussienne), dans ce cas $\sigma = 1$ et $u = 0$	28
1.9	Interconnexions neuronales.	29
1.10	Algorithme d'apprentissage supervisé.	31
1.11	Algorithme d'apprentissage non supervisé.	32
1.12	Illustration de la règle de Hebb [13]	32
1.13	Un réseau de neurone contenant deux couches [2].	35
1.14	Représentation d'une architecture d'un réseau multicouches de type feedforward [2].	38
1.15	Représentation d'un réseau récurrent.	40
1.16	Représentation d'un réseau RBF : perceptron multicouche amélioré dans lequel les neurones de la couche intermédiaire sont des neurones à base radiale.	41

1.17	Représentation du réseau de Hopfield.	43
1.18	Le réseau de Kohonen.	44
1.19	Le réseau ART de Grossberg et Carpenter.	45
3.1	(a) Thermocouple pour générer la puissance électrique (b) Thermocouple pour réfrigérer (c) Dispositif thermoélectrique contenant plusieurs thermocouples [55]	69
3.2	Modèle unidimensionnel du module Peltier	70
3.3	Illustration de l'effet Seebeck [60]	71
3.4	Illustration de l'effet Peltier [60]	72
3.5	Illustration de l'effet Thomson [60]	73
3.6	Schéma d'un module unitaire configuré pour la génération d'électricité (gauche) et réfrigération (droite) [63].	75
3.7	Le facteur de mérite de différents matériaux thermoélectriques en fonction de la température [67].	77
3.8	Modèle physique d'un élément thermoélectrique lié à deux contacts métalliques[69].	80
4.1	Modèle de réseau neuronal développé pour prédire les propriétés thermoélectriques.	95
4.2	Représentation des différentes couches du réseau de neurone artificiel	96
4.3	Erreur quadratique moyenne par rapport au nombre d'unités neuronales dans la première couche cachée.	97
4.4	Le spectre d'impédance de l'élément thermoélectrique Bi_2Te_3 . Les cercles se réfèrent aux spectres d'impédance dérivés de la réponse du réseau neuronal [98] et les lignes bleues font référence aux mesures par spectroscopie d'impédance effectuées par Jorge Garcia et al [70].	99

4.5	Résultats simulés en termes de mesures d'impédance pour l'élément thermoélectrique Bi_2Te_3 . Les cercles se réfèrent aux spectres d'impédance dérivés de la réponse du réseau neuronal et les lignes bleues font référence aux mesures par spectroscopie d'impédance [69].	100
4.6	Le spectre d'impédance de l'élément thermoélectrique Bi_2Te_3 . Les cercles se réfèrent aux spectres d'impédance dérivés de la réponse du réseau neuronal et les lignes bleues font référence aux mesures par spectroscopie d'impédance effectuées par Jorge Garcia et al [69].	100
4.7	Réponse expérimentale d'impédance (cercles) du Bi_2Te_3 de type p [69].	107
4.8	Réponse expérimentale d'impédance (cercles) du Bi_2Te_3 de type p [70].	107
4.9	L'erreur quadratique moyenne par rapport au nombre d'itérations du modèle neuronal.	110

Liste des tableaux

4.1	Les valeurs des différents paramètres thermoélectriques de l'élément Bi_2Te_3 caractérisé par une surface de $1.4 \times 1.4 \text{ mm}^2$ et 1.6 mm de longueur.	105
4.2	Les valeurs des différents paramètres thermoélectrique de l'élément Bi_2Te_3 caractérisé par une surface de $1.4 \times 1.4 \text{ mm}^2$ et 1.6 mm de longueur.	106
4.3	Les valeurs des différents paramètres thermoélectriques de l'élément Bi_2Te_3 caractérisé par une surface de 1 mm^2 , et 1.5mm de longueur.	106
4.4	Les valeurs du coefficient de Seebeck S, de la conductivité thermique de la diffusivité thermique et du coefficient de Peltier de quelques éléments.	109

Résumé

L'objectif de notre contribution porte sur l'utilisation des Réseaux de Neurones Artificiels pour la caractérisation des propriétés des éléments thermoélectriques. A cause de la nature fortement non linéaire des phénomènes thermoélectriques, cette approche présente plusieurs avantages, d'une part la possibilité de tenir compte de plusieurs paramètres en même temps et d'autre part la possibilité d'étudier l'intégralité de ces propriétés. Dans une telle approche neuronale, une description adéquate des matériaux étudiés est développée par la construction d'un réseau de neurones artificiels capable d'optimiser les données et de calculer les paramètres thermoélectriques. Pour cela, nous définissons l'impédance électrique du système dans des conditions adiabatiques en résolvant l'équation de la chaleur sur le domaine de fréquences choisies. Nos résultats ont été validés en les comparant avec ceux obtenus par la technique de caractérisation par spectroscopie d'impédance et avec les données expérimentales électrochimiques. Les valeurs obtenues à l'aide de l'approche neuronale ont permis de définir les différentes propriétés caractérisant ces éléments. Une validation des résultats obtenus avec les valeurs données dans la littérature permet de confirmer la fiabilité et la robustesse du modèle de réseau de neurones choisi.

Abstract

The objective of our contribution relates to the use of an Artificial Neural Network for the characterization of the thermoelectric elements properties, because of the nonlinear nature of thermoelectric phenomena, this approach has several advantages, on the one hand the possibility to take into account several parameters at the same time, on the other hand the possibility of studying all of these properties. In such a neuronal approach, an adequate description of the studied

materials is developed by the construction of an artificial neuron network capable of optimizing data and calculating thermoelectric parameters. For this, we define the electrical impedance of the thermoelectric system under adiabatic conditions by solving the heat equation over the chosen frequency domain. Our results were validated by comparing them with those obtained by the technique of characterization by impedance spectroscopy and with the electrochemical experimental data. The values obtained using the neuronal approach made it possible to define the different properties characterizing these elements. A validation of the results obtained with the values given in the literature makes it possible to confirm the reliability and the robustness of the chosen neural network.

الهدف من مساهمتنا يتعلق باستخدام الشبكات العصبونية الاصطناعية من أجل توصيف الخصائص الكهرو حرارية لمختلف العناصر. بما أن الظواهر المصاحبة لعملية التحول الكهرو حراري ذات طبيعة لا خطية يعتبر هذا النهج مفيدا لما له من مزايا تسمح بالتعامل مع معطيات متنوعة في آن واحد من ناحية ومن ناحية أخرى إمكانية تحليل ودراسة واستخراج مجمل الخصائص التي تميز عملية التحول الكهرو حراري. سيتم تطوير وصف مناسب للعناصر المدروسة من خلال بناء شبكة عصبونية اصطناعية قادرة على تحسين البيانات وحساب قيم خواص المواد الكهرو حرارية وذلك بتوظيف معادلة المعاوقة الكهربائية التي نتحصل عليها من خلال حل معادلة الحرارة في مجال الترددات. تم التحقق من صحة النتائج المتحصل عليها من خلال مقارنتها مع تلك التي تم الحصول عليها بطريقة التوصيف بالتحليل الطيفي للمعاوقة ومع البيانات التجريبية المتوفرة. أثبتت المقارنة أن النتائج التي تم الحصول عليها باستخدام الشبكات العصبونية متطابقة مع مختلف القيم المتوفرة مما يؤكد موثوقية وقدرة النموذج العصبوني المختار على تحديد مختلف الخصائص التي تميز العناصر الكهرو حرارية.

Introduction Générale

Pour comprendre les réseaux de neurones, il ne suffit pas de s'intéresser aux mathématiques et à la technologie. Il s'agit de remonter à un élément essentiel de la vie, l'être humain. L'homme a en effet toujours cherché à reproduire les différentes capacités qu'il utilise au quotidien en simulant le cerveau humain par des procédés informatiques réalisés par les réseaux de neurones artificiels. Simuler l'intelligence humaine c'est modéliser mathématiquement le cerveau afin de reproduire certaines caractéristiques comme la capacité de mémoire c'est à dire se souvenir de ce qui s'est passé, la capacité d'apprentissage par des algorithmes qui sont capables d'apprendre de leur environnement même s'il évolue et la capacité à traiter des informations incomplètes.

La caractéristique la plus importante du réseau neuronal est le traitement parallèle des informations par tous les liens. Avec une énorme quantité de connexions inter-neuronales, il est possible d'accélérer considérablement le processus de traitement des informations. De plus, la propriété fondamentale des réseaux de neurones, l'approximation universelle parcimonieuse, fait de ceux-ci une représentation mathématique très avantageuse pour les modélisations statique et dynamique non linéaires des phénomènes. En effet, il a été démontré qu'un réseau de neurones de type perceptron, ayant une couche cachée est un " approximation universel parcimonieux ", c'est à dire qu'il peut approcher toute fonction multi-variables non linéaires et ses dérivées avec une précision à priori fixée, pourvu qu'il soit soumis

à un apprentissage efficace.

La demande mondiale croissante en matière de production, de conservation et de gestion de l'énergie nous oriente vers des énergies alternatives respectueuses de l'environnement. Les sources d'énergie renouvelable actuelles (solaire, hydraulique et géothermique) ne sont pas suffisantes, principalement à cause de leur coût relativement élevé, alors qu'il existe des ressources abondantes, à faible coût et encore relativement peu exploitées sous forme de chaleur. Il s'agit bien évidemment de toute la chaleur résiduelle produite par de nombreuses et diverses sources (chauffage domestique, échappement de voitures, industrie,...) qui ne peut pas être exploitée efficacement dans la conversion de la chaleur perdue en électricité.

Les dispositifs thermoélectriques peuvent convertir directement la chaleur perdue en électricité. La partie active de ces dispositifs est constituée de matériaux semi-conducteurs qui pourraient contribuer à réduire notre dépendance aux combustibles fossiles et les problèmes environnementaux associés, ils ont considérablement contribué au développement de certaines missions spatiales en fournissant l'énergie électrique nécessaire au bon fonctionnement des émetteurs de signaux radio, des ordinateurs de bord, des gyroscopes et des systèmes de navigation de plusieurs sondes spatiales lancées dans l'espace lointain. Dans la perspective d'applications à plus grande échelle, ils sont intéressants pour la climatisation ou la réfrigération.

La connaissance des propriétés thermiques des matériaux permet de modéliser les transferts thermiques dans les systèmes complexes et de prédire leur comportement. Il existe différentes techniques pour caractériser les propriétés thermoélectriques. La conductivité thermique, la conductivité électrique et le coefficient Seebeck sont les trois grandeurs recherchées pour le calcul du facteur de mérite qui détermine les performances et l'efficacité d'un module thermoélectrique, sou-

vent la caractérisation du matériel implique l'utilisation jusqu'à trois équipements différents. En particulier, la conductivité thermique est difficile à mesurer et nécessite un appareillage très coûteux. C'est pour cela qu'il vaut la peine de développer de nouvelles méthodes et techniques de caractérisation.

Les réseaux de neurones sont un outil puissant capable de résoudre une variété de tâches : l'approximation des fonctions, la reconnaissance des caractères et la modélisation des problèmes physiques. Leur puissance est utilisée pour développer des solutions alternatives de plus en plus efficaces. Nous allons utiliser l'approche de l'apprentissage par réseau de neurone pour modéliser les caractéristiques de l'impédance totale obtenue en résolvant l'équation de la chaleur d'un élément thermoélectrique de type n ou de type p. Dans une telle approche neuronale, une description adéquate des éléments étudiés sera prouvée par la possibilité de construire un réseau de neurones artificiels capables d'optimiser les données et de calculer les paramètres thermoélectriques à savoir : la conductivité thermique, la diffusivité thermique, le coefficient Seebeck, la résistance thermique et la fréquence caractéristique de diffusion thermique estimées par l'analyse de l'impédance équivalente de l'interface étudié, nos résultats ont également été validés en les comparant avec ceux obtenus par caractérisation par spectroscopie d'impédance et avec les données expérimentales électrochimiques. Les valeurs extraites de l'approche neuronale ont permis de définir les différentes propriétés caractérisant ces éléments. Une formation et une validation réussies ont prouvé la crédibilité de l'approche de réseau de neurones où les quantités de sortie calculées par le réseau étaient en excellent accord avec les valeurs de données qui existent. Le modèle neuronal peut être considéré comme étant de nature universelle puisqu'il peut être étendu en tant que nouvelle méthode d'approximation utilisée dans des systèmes complexes, en particulier dans la mesure où il est utile sans aucune contrainte quant au nombre de composants pouvant être étudiés.

Nous expliquerons les principes fondamentaux qui justifient l'intérêt pratique des réseaux de neurones, et nous situerons ces derniers dans la perspective des méthodes classiques de traitement statistique de données ; nous montrerons que la technique des réseaux de neurones formels doit être considérés comme une extension puissante de techniques bien connues des ingénieurs, telles que la régression, et ceci en développant une technique de caractérisation capable d'extraire certaines quantités physiques de certains matériaux. En effet, l'application des réseaux de neurones à la physique est un sujet qui se développe rapidement, on essaye de prédire l'évolution d'un système par la résolution des équations fondamentales. De plus, on prédit un phénomène à partir d'une base de données d'exemples en capturant des formes de régularité de cette évolution, des résultats de plus en plus précis sont obtenus avec les réseaux de neurones en physique des matériaux , c'est dans ce contexte que nous allons développer une étude théorique d'un modèle de réseau de neurones pour définir et prédire les différents paramètres caractérisant les matériaux thermoélectriques. La topologie du modèle de réseau neuronal est fortement corrélée au problème étudié et aux processus sélectionnés pour traiter complètement les données, ce qui a donné lieu à de nouvelles découvertes dans le domaine de la thermoélectricité, s'inscrivant parfaitement dans la recherche de nouvelles sources d'énergie respectueuses de l'environnement, ceci a suscité un intérêt grandissant au sein de la communauté scientifique.

Dans ce travail le premier chapitre est consacré à la présentation de la théorie des réseaux de neurones en détaillant les différents éléments de base de cet outil mathématique, après avoir exposé un aperçu sur les différentes règles d'apprentissage, nous citerons les différents modèles à base de réseaux de neurone à savoir les réseaux à interconnexion complète et les réseaux à propagation avant. Nous limitons notre présentation à des aspects qui seront utiles pour la suite de ce travail.

Le deuxième chapitre est dédié à la présentation des méthodes d'apprentissage

qui s'appuient sur la rétropropagation de l'erreur. Il s'agit en fait de trouver les paramètres du modèle qui minimisent l'erreur. L'apprentissage étant un problème d'optimisation non linéaire, il est généralement résolu par des méthodes basées sur le gradient. Dans un souci de clarté et afin de faciliter la compréhension, il nous a paru nécessaire de distinguer, dans ce chapitre, l'étape de calcul du gradient de la fonction de coût par rapport aux paramètres (rétropropagation du gradient de l'erreur) de l'étape de modification des paramètres (méthodes du gradient simple, de Newton et de Levenberg Marquardt). Nous terminons le chapitre par un énoncé des propriétés spécifiques aux réseaux de neurones, suivi d'une présentation de l'intérêt de ces techniques neuronales. Cette partie s'achève par une revue générale des domaines d'applications.

Dans le troisième chapitre nous aborderons le domaine de la thermoélectricité, après un bref aperçu sur les phénomènes thermoélectriques, nous introduirons les principes de base de la conversion d'énergie par effet thermoélectrique et les principaux paramètres à contrôler pour l'obtention d'un bon rendement de conversion. Nous discuterons ensuite des performances thermoélectriques d'un dispositif idéal afin de montrer que le facteur de mérite adimensionnel zT représente le paramètre essentiel permettant de comparer les différents composés thermoélectriques étudiés. Par la suite nous allons développer le calcul de l'impédance électrique d'un élément thermoélectrique dans les conditions adiabatiques qui peut être simulée par spectroscopie d'impédance en un circuit équivalent.

Le quatrième chapitre présente la méthodologie suivie afin d'atteindre notre objectif de modélisation pour extraire et prédire les propriétés thermoélectriques des matériaux. Nous appliquons les étapes de conception de réseaux de neurones décrites au premier et au deuxième chapitre, à savoir l'élaboration des bases d'apprentissage et de test et le choix de la structure neuronale à utiliser pour obtenir de bonnes prédictions, le modèle neuronal est retenu dans des conditions

données (température, forme) et les unités d'entrée du réseau sont définis par l'impédance dans les plans réels et imaginaires, la collectes de données est extraite en définissant chaque paramètre dans un intervalle par rapport à ces valeurs empiriques et les unités de sortie incluent les différentes caractéristiques des matériaux. En raison de la convergence locale liée à chaque condition aux limites, un ensemble aléatoire de poids initiaux et de valeurs de biais a été développé, de plus, la phase d'entraînement du réseau neuronal est réalisée en utilisant l'algorithme de Levenberg Marquardt. Pour l'étape de simulation, nous révélons principalement la capacité du modèle neuronal à prédire le coefficient Seebeck, la conductivité thermique, la diffusivité thermique et le coefficient Peltier de certains éléments. La dernière partie est réservée aux résultats obtenus en vue de la validation du modèle développé et à la comparaison de ces résultats avec ceux obtenus par caractérisation par spectroscopie d'impédance.

Ce manuscrit se termine par une conclusion générale qui rappelle les résultats originaux obtenus lors de ce travail. Nous aborderons enfin les perspectives futures à explorer à l'aide de la méthode des réseaux de neurones.

1 Chapitre 1: Généralités sur les réseaux de neurones

1.1 Introduction

En lisant ces mots, vous utilisez un réseau neuronal biologique complexe. Vous disposez d'un ensemble hautement inter-connecté de quelques 10^4 neurones pour faciliter votre développement. A travers l'histoire, les êtres humains ont toujours eu conscience de la place particulière qu'ils occupent dans le monde du vivant : empathie, humour, créativité, imagination, conscience, pensée, philosophie, littérature, culture, art, science, histoire, politique, invention, progrès, commerce, éducation, sont autant de termes qui les différencient des animaux et des insectes. Mais ces termes ont un autre point commun. Ils sont tous reliés d'une certaine manière à une propriété qui pourrait figurer en tête de cette précédente liste de l'intelligence [1].

Les sciences liées à la biologie considèrent naturellement que l'intelligence est liée au fait de posséder des neurones. Les scientifiques ne font que commencer à comprendre le fonctionnement des réseaux neuronaux biologiques. Il est généralement admis que toutes les fonctions neuronales biologiques, y compris la mémoire, sont stockées dans les neurones et leurs connexions. L'apprentissage est considéré comme l'établissement de nouveaux liens entre ces neurones ou la modification des connexions existantes. Le système nerveux central humain est un système de traitement de l'information stupéfiant en termes de capacités, de polyvalence, d'adaptabilité et de faible consommation d'énergie. Sa structure complexe se compose de milliards de neurones, inter-connectés par plusieurs trillions de synapses, formant des grappes spécialisées.

Les neurones que nous considérons ne sont pas biologiques. Ils sont extrêmement de simples abstractions de neurones biologiques, réalisées sous forme d'éléments dans un programme. Les réseaux de neurones artificiels n'ont pas une fraction du pouvoir du cerveau humain, mais ils peuvent être formés pour exécuter des fonctions utiles. Un réseau de neurones artificiels (RNA) est constitué par l'interconnexion d'unités élémentaires de traitement de l'information, dont le principe de fon-

tionnement est calqué sur celui des neurones biologiques. Un réseau possède des capacités de stockage et de traitement de l'information, imputables aux poids des connexions entre chaque neurone. Ces poids sont calculés par un processus d'adaptation également appelé apprentissage.

1.2 Inspiration biologique

Les neurones biologiques, aussi appelés cellules nerveuses, sont les principaux éléments constitutifs du système nerveux central humain. Ce dernier se compose principalement du cerveau et du système nerveux périphérique constitué des nerfs qui traversent le corps humain, la plupart par la moelle épinière. Un neurone biologique est une cellule vivante consacrée au traitement de l'information et il comporte généralement trois zones de cellules.

- Le corps cellulaire, ou soma, constitué du cytoplasme et d'un noyau contenant les gènes.
- L'arbre dendritique, collectant les signaux provenant d'autres cellules nerveuses (ou de l'extérieur).
- L'axone, diffusant le signal et se divisant normalement à son extrémité afin d'entrer en contact.

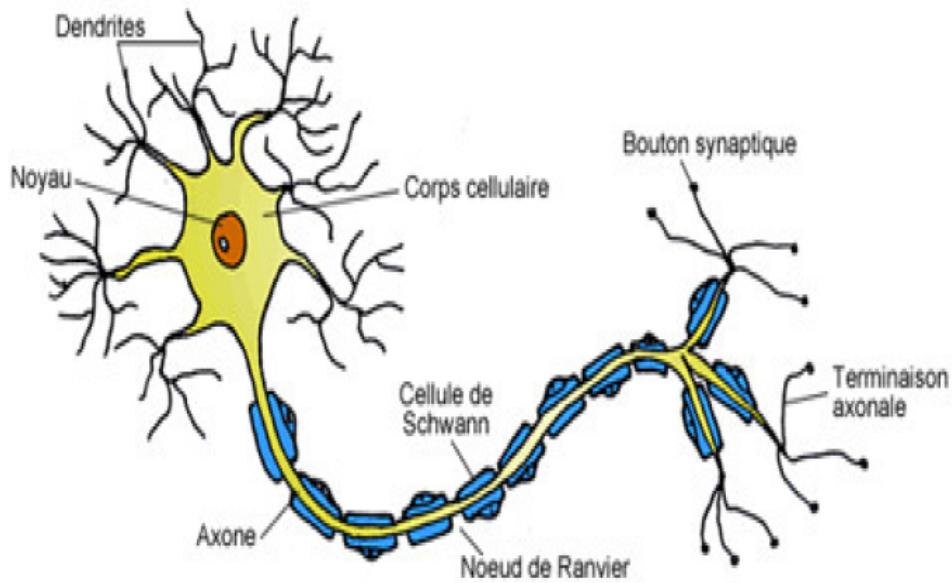


Figure 1.1 – Le neurone biologique [2]

Un processus électrochimique complexe permet la transmission de signaux entre les différents neurones. Le signal, qui est reçu à travers les dendrites (les récepteurs), est un potentiel d'action, c'est-à-dire une activation électrique d'une durée de 1 à 5 ms et d'une amplitude d'environ 100 mV, qui se propage le long de l'axone. Arrivé à l'extrémité de l'axone (l'émetteur), le potentiel d'action engendre un signal de nature chimique sous forme d'un neurotransmetteur diffusé dans la fente synaptique. Une synapse est une structure élémentaire et une unité fonctionnelle entre deux neurones (un brin d'axone d'un neurone et une dendrite d'un autre). Lorsque le neurotransmetteur atteint le neurone suivant, il génère un potentiel synaptique. Ce dernier provoque une excitation ou une inhibition du neurone. Le neurone reçoit de tels potentiels synaptiques de toutes les cellules nerveuses avec lesquelles il est en liaison. Il effectue une sommation de ces informations. Si le résultat obtenu est supérieur à un seuil, le soma génère un potentiel d'action d'amplitude et de durée fixes. Dans les détails il y a beaucoup de varia-

tions de ce concept. Pour la suite, le simple concept d'une sommation d'entrées suivie d'un effet seuil est suffisant.

1.3 Historique du neurone artificiel (ou neurone formel)

L'histoire des réseaux de neurones artificiels est tissée à travers des découvertes conceptuelles et des développements technologiques survenus à diverses époques. L'histoire a commencé par:

- la présentation des neurologues Warren Sturgis Mc Culloch et Walter Pitts d'un modèle simplifié du neurone biologique communément appelé neurone formel [3] en 1943. Ils montrèrent que des réseaux de neurones formels relativement simples sont capables de réaliser des fonctions logiques et arithmétiques. Cependant, leurs travaux n'ont pas donné d'indication sur une quelconque méthode permettant d'estimer les coefficients synaptiques.
- En 1949, le psychologue Donald Hebb a introduit le terme connexionnisme pour parler des modèles massivement parallèles et connectés. Il a proposé de nombreuses règles de mise à jour des coefficients synaptiques dont la plus connue à cette époque est la "règle de Hebb" [4]. De nombreux modèles de réseaux aujourd'hui s'inspirent encore de la règle de Hebb.
- En 1958, le psychologue Frank Rosenblatt a introduit le modèle du perceptron [5]. c'est un modèle conçu pour différencier des formes simples et calculer certaines fonctions logiques. Il est inspiré du système visuel et il a réussi à l'appliquer pour la reconnaissance des formes.
- En 1982, Hopfield a fini par démontrer l'intérêt des réseaux entièrement connectés aboutissant au Modèle de Hopfield [6]. Ce modèle discute de la théorie du fonctionnement et des possibilités des réseaux de neurones. Parallèlement, Werbos a conçu un mécanisme d'apprentissage pour les

réseaux multicouches de type perceptron [7] : la rétropropagation (Back-Propagation) [8]. Dès lors, le domaine des réseaux de neurones artificiels a connu un essor considérable et a suscité un intérêt toujours croissant. Il est devenu un domaine où bouillonnent constamment de nouvelles théories, de nouvelles structures et de nouveaux algorithmes. Ils servent aujourd'hui à toutes sortes d'applications dans divers domaines. Ils sont utilisés dans le traitement des signaux, la reconnaissance des formes, l'approximation des fonctions, la prédiction, l'optimisation et la mémoire associative, le diagnostic médical, en robotique,...etc

1.4 Le neurone formel de McCulloch et Pitts

Le neurone formel schématisé dans la figure (1.2) est une modélisation mathématique du neurone biologique : Le neurone formel d'un point de vue mathéma-

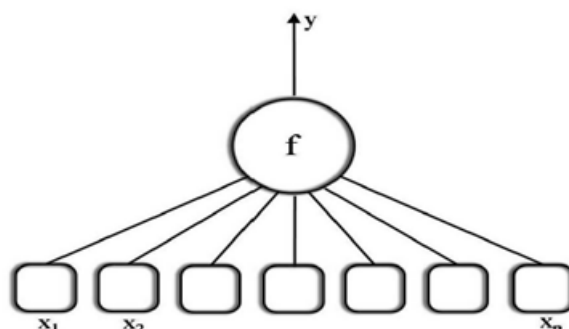


Figure 1.2 – Schématisation algorithmique d'un neurone artificiel [9].

tique est une fonction algébrique non linéaire, paramétrée à valeurs bornées qui effectue la somme pondérée de ses entrées x_i . A chacune de ces entrées est associé un paramètre appelé coefficient ou poids de connexions w_i qui caractérise la force des liens dans les connexions. Le résultat de cette somme est ensuite transformée par une fonction de transfert ϕ appelée fonction d'activation c'est une schématisation du fonctionnement du neurone qui sert à activer ou neutraliser le signal

qu'elle reçoit et qui produit la sortie y du neurone. Chaque neurone est doté d'une sortie unique y tel que :

$$y = \phi\left(\sum_{i=1}^n w_i x_i + b\right) \quad (1.1)$$

b est le biais ou autrement dit : seuil d'activation du neurone formel, cette représentation est le reflet de l'inspiration biologique en effet :

- Les synapses sont modélisées par des poids
- Le soma ou corps cellulaire est modélisé par la fonction de transfert
- L'axone par l'élément de sortie

1.5 Les fonctions d'activation

La fonction d'activation opère une transformation d'une combinaison fine des signaux d'entrée, cette fonction peut aussi porter le nom de fonction de combinaison ou fonction de seuillage son concept vient de l'idée de mimer le fonctionnement du potentiel d'action d'un neurone biologique, son choix se révèle être un élément constitutif important des réseaux de neurones. En effet, les propriétés de la fonction d'activation influent sur celle du neurone formel et il est donc important de bien choisir celle-ci pour obtenir un modèle utile en pratique. Il existe plusieurs types de fonctions d'activation qu'on pourrait utiliser selon les applications souhaitées. A titre illustratif, voici quelques fonctions couramment utilisées.

1.5.1 Fonction seuil

Selon la propriété du neurone biologique, chaque neurone possède un seuil d'excitation. Ce seuil est toujours binaire, soit ON ou OFF. Cela signifie que lorsque la stimulation des signaux d'entrée est plus grande qu'un seuil, il y aura une réponse de sortie (ON). Sinon, aucune réponse n'est donnée à partir de ce neurone (OFF). Inspirée par la propriété du neurone biologique, la fonction d'activation

seuil est définie mathématiquement comme :

$$\phi(v) = \begin{cases} 0 & \text{si } v < 0, \\ 1 & \text{si } v \geq 0, \end{cases} \quad (1.2)$$

En pratique, la fonction d'activation du neurone artificiel doit être plus générale

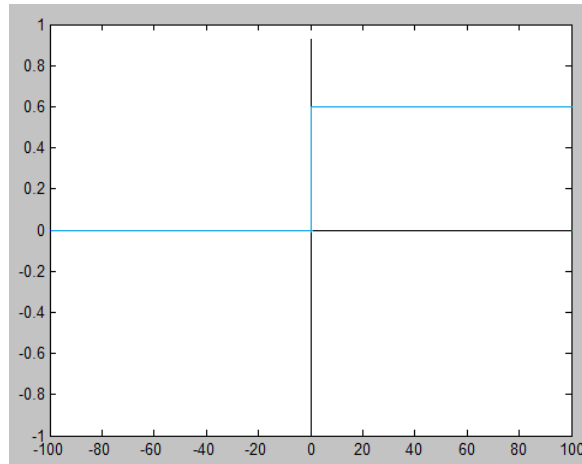


Figure 1.3 – Représentation de la fonction seuil.

que la simple propriété binaire, c'est-à-dire continue et dérivable. Pour cette raison, plusieurs types de fonctions d'activation sont développés. Mais le choix de cette dernière, dans le domaine des réseaux de neurones, doit être précis pour arriver à un modèle efficace en pratique.

1.5.2 Fonction linéaire

La fonction linéaire est la fonction d'activation la plus fondamentale. Elle agit comme une amplification linéaire des signaux d'entrée du réseau :

$$\phi(v) = kv, \quad (1.3)$$

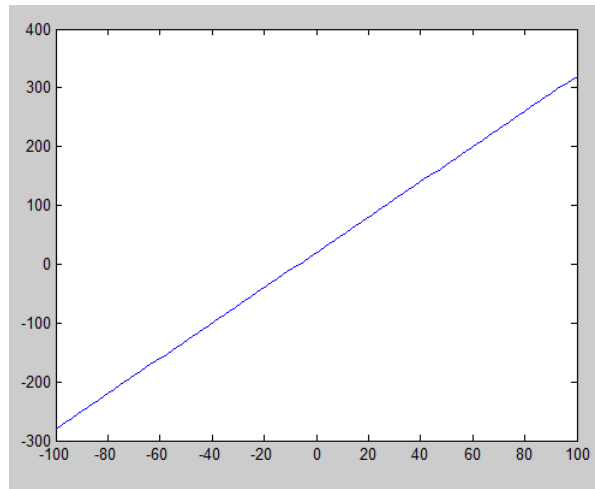
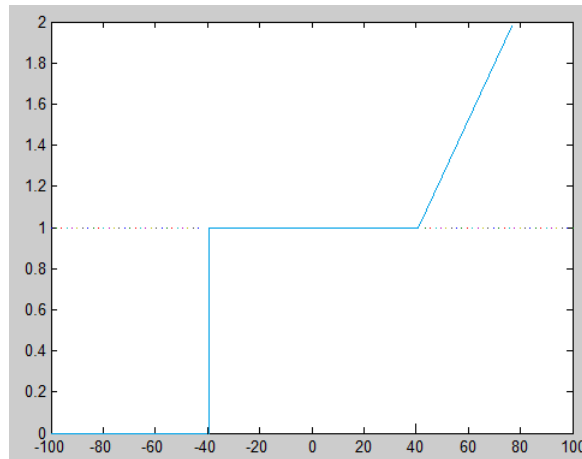


Figure 1.4 – Représentation de la fonction linéaire.

1.5.3 Fonction linéaire par morceaux (Piecewise)

Bien que la fonction linéaire soit très simple à utiliser, elle limite la performance du réseau et diminue la non-linéarité. Pour éviter cet inconvénient, on introduit une fonction linéaire par morceaux :

$$\phi(v) = \begin{cases} kv & \text{si } v \geq \theta, \\ 1 & \text{si } -\theta \leq v \leq \theta \\ 0 & \text{si } v \leq -\theta, \end{cases} \quad \text{ou } \theta \text{ est un seuil,} \quad (1.4)$$

Figure 1.5 – Représentation de la fonction linéaire par morceaux ($\theta = 40$).

1.5.4 Fonction sigmoïde

La fonction sigmoïde est la fonction la plus couramment utilisée dans le réseau de neurones en raison de sa capacité différentiable. Deux types de fonctions sigmoïdes sont habituellement utilisées : la fonction logistique représentée dans la figure (1.6)

$$\phi(v) = \frac{1}{1 + e^{-v}}, \quad (1.5)$$

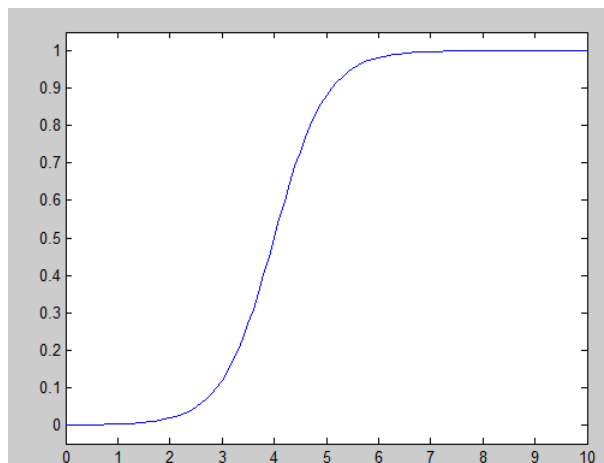


Figure 1.6 – Représentation de la fonction sigmoïde logistique (logsig).

Et la fonction hyperbolique représentée dans la figure (1.7).

$$\phi(v) = \frac{1 + e^{-v}}{1 - e^{-v}}, \quad (1.6)$$

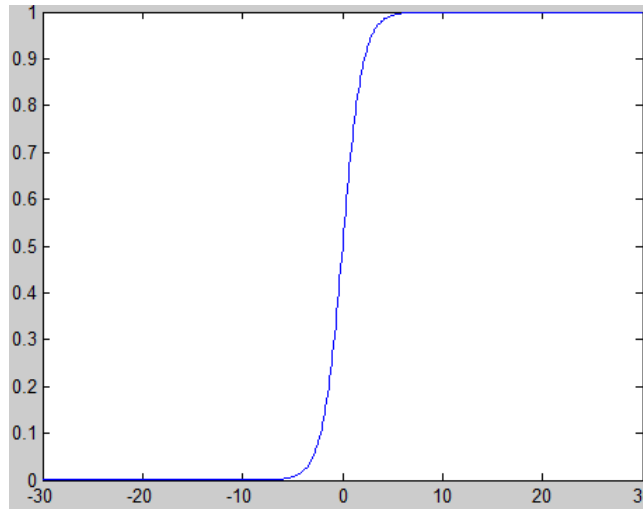


Figure 1.7 – Représentation de la fonction sigmoïde tangente hyperbolique (tansig).

1.5.5 Fonction radiale

Introduite et utilisée dans le cadre de l'interpolation.

$$\phi(v, u, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{v - u}{\sigma}\right)^2\right], \quad (1.7)$$

Avec σ la coordonnée du centre de la gaussienne et u son écart type.

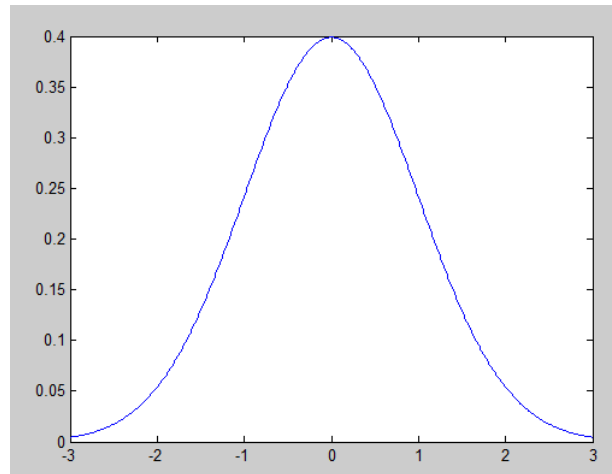


Figure 1.8 – Représentation de la fonction radiale (gaussienne), dans ce cas $\sigma = 1$ et $u = 0$.

1.6 Interconnexions neuronales

Un réseau de neurones est constitué d'un ensemble d'unités (neurones formels) interconnectées. Chaque neurone possède des entrées, des sorties et des paramètres, il est capable de calculer ses sorties en fonction de ses entrées et de ses paramètres par l'application de la fonction d'activation ϕ qui est fixée pour chaque neurone artificiel selon des objectifs de la modélisation. Les connexions que l'on appelle aussi synapses transportent les sorties de certains neurones vers les entrées d'autres neurones, ces neurones sont organisés en couches :

- La couche d'entrée est la première couche du réseau. Elle contient les variables d'entrées qui peuvent être les données sources que l'on veut utiliser pour l'analyse.
- Les couches cachées. Le choix de leur taille n'est pas implicite et doit être ajusté.
- La couche de sortie donne directement le résultat obtenu par le réseau après compilation. Sa taille est déterminée par le nombre de variables que l'on veut en sortie. L'application de la fonction d'activation ϕ qui est fixée

pour chaque neurone artificiel selon des objectifs de la modélisation, afin de donner un nouveau seuil d'activation. Ces étapes se propagent entre les neurones d'entrée jusqu'à arriver aux neurones de sorties qui permettent de reproduire le comportement du système à modéliser. Les connexions entre les neurones artificiels dans un réseau de neurones artificiels déterminent sa topologie.

L'application de la fonction d'activation ϕ qui est fixée pour chaque neurone artificiel selon des objectifs de modélisation afin de donner un nouveau seuil d'activation. Ces étapes se propagent entre les neurones d'entrée jusqu'à arriver aux neurones de sorties qui permettent de reproduire le comportement du système à modéliser. Les connexions entre les neurones artificiels dans un réseau de neurones artificiels déterminent sa topologie.

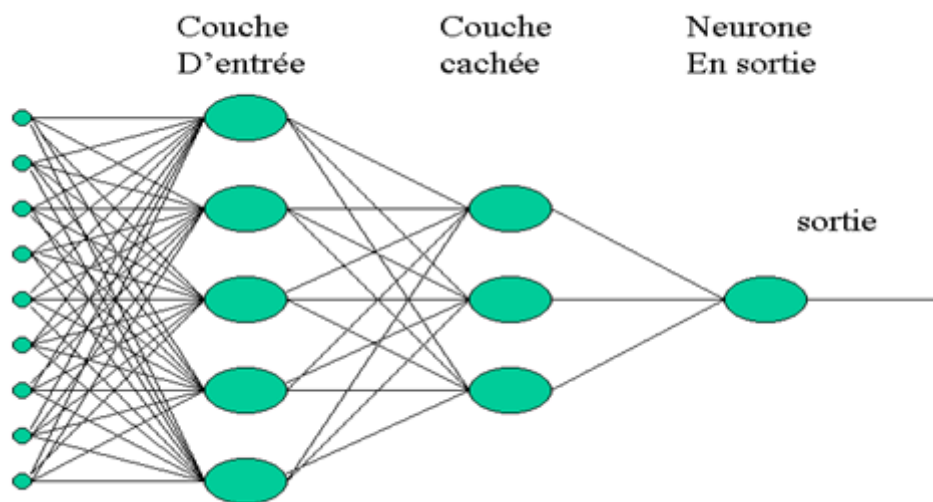


Figure 1.9 – Interconnexions neuronales.

1.7 Apprentissage dans les réseaux de neurones

Parmi les propriétés désirables pour un réseau de neurones, la plus fondamentale est sûrement sa capacité d'apprendre de son environnement et d'améliorer ainsi sa performance à travers le processus d'apprentissage. Ce dernier est tout simplement une phase de développement du réseau de neurones durant laquelle le comportement du réseau est modifié jusqu'à l'obtention du comportement désiré [10]. Au cours de cette phase, les paramètres des neurones sont alors estimés afin que le réseau remplisse au mieux la tâche qui lui est affectée. On distingue deux types d'apprentissage : l'apprentissage (supervisé) et l'apprentissage (non supervisé) [11]. Cette distinction repose sur la forme des exemples d'apprentissage. Dans le cas de l'apprentissage supervisé, les exemples sont des couples (Entrée, Sortie associée) alors que l'on ne dispose que des valeurs (Entrée) pour l'apprentissage non supervisé.

1.7.1 Apprentissage supervisé

Lorsque l'on connaît en tout point ou en un ensemble de points les valeurs que doit avoir la sortie du réseau en fonction des entrées correspondantes, on parle d'apprentissage supervisé. Pour les réseaux à apprentissage supervisé, on présente donc au réseau des entrées et au même temps les sorties désirées. Le réseau doit alors se reconfigurer, c'est à dire calculer ses poids afin que la sortie qu'il donne corresponde bien à la sortie souhaitée. Ce type d'apprentissage est utilisé pour des tâches d'approximation, de modélisation statique et dynamique, de classification, etc.

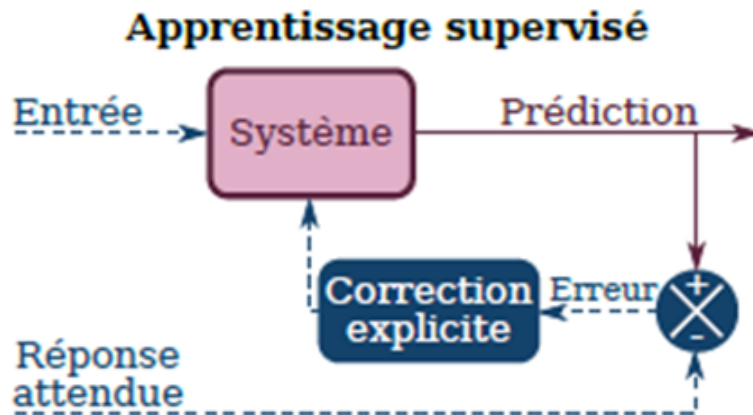


Figure 1.10 – Algorithme d'apprentissage supervisé.

1.7.2 Apprentissage non supervisé

Ce type d'apprentissage est utilisé dans un but de visualisation ou d'analyse de données. En effet, lorsqu'on dispose d'un ensemble de données et l'on cherche à les regrouper selon des critères de ressemblance qui sont inconnus à priori, on parle d'apprentissage non supervisé. Dans ce cas, on présente une entrée au réseau et on le laisse évoluer librement jusqu'à ce qu'il se stabilise. D'une manière générale, cette étape, appelée phase de développement du réseau de neurones, requiert l'utilisation de techniques d'optimisation qui consistent à déterminer les paramètres optimaux par minimisation directe d'une fonction, appelée fonction de coût, vis-à-vis des paramètres du modèle. La procédure adaptative de mise à jour des poids obéit essentiellement à un objectif ; la recherche du minimum de la fonction de coût (ou fonction d'erreur) dans un espace multidimensionnel [12].

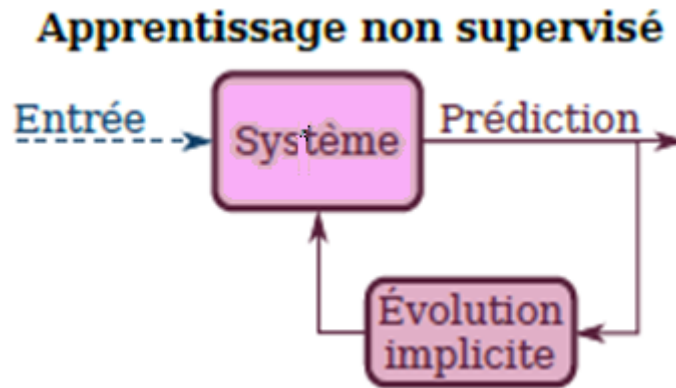


Figure 1.11 – Algorithme d'apprentissage non supervisé.

1.7.3 Les règles d'apprentissage

Loi de Hebb

C'est une règle qui s'inspire des travaux du neurophysiologiste Donald Hebb [4], elle stipule que si deux neurones sont activés dans le même temps (1.12) alors

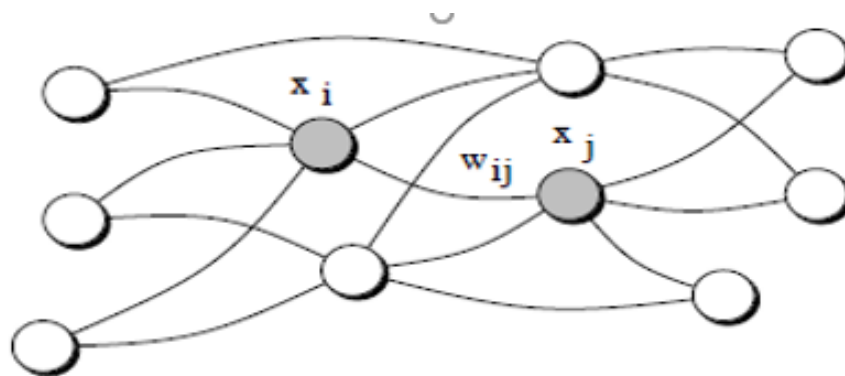


Figure 1.12 – Illustration de la règle de Hebb [13]

le poids de la connexion qui les relie est augmenté. On évalue cette augmentation comme :

$$w_{ij} = R \cdot x_i \cdot x_j, \quad (1.8)$$

Où R est une constante d'apprentissage, x_i la valeur de l'activation du neurone i et x_j la valeur d'activation du neurone j autrement dit, il y a renforcement de la connexion entre i et j lorsque i et j sont activés ($x_i > 0$ et $x_j > 0$).

Règle de Widrow-Hoff

La règle de Widrow-Hoff [14]. reprend le principe de la loi de Hebb mais en y ajoutant une quantification de l'erreur de sortie de chaque neurone. Il s'agit par conséquent d'une règle d'apprentissage supervisée. Pour chaque neurone j cette règle met en oeuvre une descente de gradient jusqu'à l'état désiré d_j du neurone j . Les connexions peuvent être renforcées ou affaiblies. L'incrément w_{ij} est donné par :

$$w_{ij} = R.(d_j - x_j)x_i, \quad (1.9)$$

Où R est une constante d'apprentissage similaire à celle utilisée par la loi de Hebb, x_i la valeur d'activation du neurone i , d_j la valeur d'activation désirée du neurone j et x_j la valeur de l'activation du neurone j . L'apprentissage prend donc en compte l'écart entre l'activation réelle et l'activation désirée du neurone postsynaptique $j(d_j, x_j)$ mais il y a apprentissage uniquement lorsque i est activé et participe bien à l'activation ou non de $j(x_i > 0)$. C'est encore aujourd'hui la loi dédiée aux perceptrons.

Règle de Grossberg

Enfin, une dernière règle d'apprentissage se démarque clairement des deux précédentes : la loi de Grossberg [15]. Cette loi est non supervisée. Son principe est d'appliquer le renforcement sur toutes les connexions afférentes à un seul neurone, qu'on appelle le neurone gagnant, dans le but de favoriser l'activation future du même neurone dans des conditions équivalentes.

$$w_{ij} = R.x_i.(x_j w_{ij}), \quad (1.10)$$

La loi de Grossberg s'applique uniquement à des topologies à deux couches. Lors de l'application de la règle (1.10), le neurone a_i qui représente le neurone gagnant est le seul activé sur sa couche. Les connexions entre chaque neurone de la seconde couche et l'unité gagnante x_i de la première couche sont renforcées proportionnellement à leur poids actuel. C'est la règle d'apprentissage utilisée dans les réseaux ART de Carpenter et Grossberg ou dans les cartes auto-organisatrices de Kohonen [16].

1.8 Modèles à base de réseaux de neurones

Initialement étudiés en vue de modéliser le comportement du cerveau humain, les modèles à base de réseaux de neurones sont aujourd'hui des outils de calculs mathématiques sophistiqués utilisés dans des domaines très divers. Il existe différents types de réseaux de neurones artificiels pour réaliser différentes tâches :

1.8.1 Les réseaux à propagation avant

Les réseaux à propagation avant ou feedforward networks [17] sont des réseaux en couche dans lesquels l'information se propage uniquement des couches d'entrée vers les couches de sortie.

a-Le perceptron simple

Le perceptron simple ou mono-couche de Frank Rosenblatt est le premier réseau de neurones fonctionnel muni d'une méthode d'apprentissage, c'est un système qui combine le neurone formel de McCulloch et Pitts avec une règle d'apprentissage. Il s'agit d'un classifieur linéaire à une couche et une sortie (binaire également), permettant de séparer en deux classes les éléments présentés en entrée [18].

La figure (1.13) montre un perceptron monocouche, avec n entrées, une couche contenant h_1 unités de neurones et m neurones dans la couche de sortie. Les

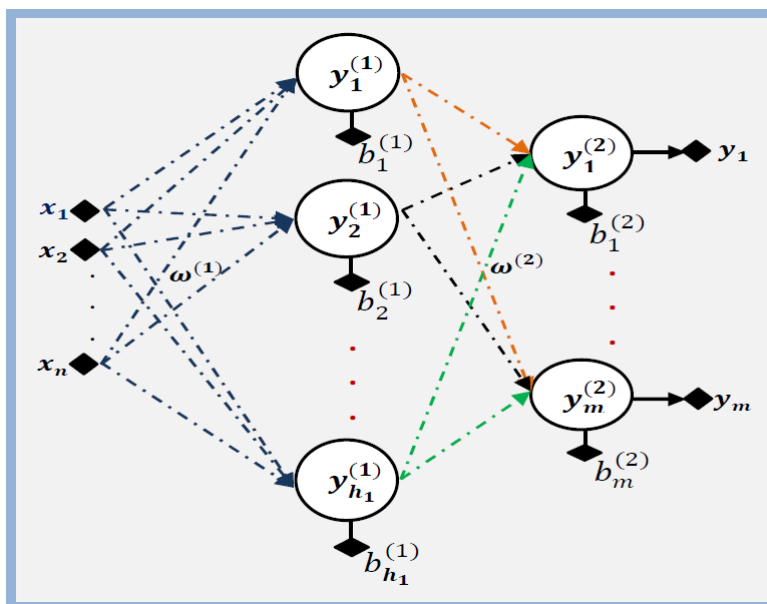


Figure 1.13 – Un réseau de neurone contenant deux couches [2].

paramètres, dans les deux couches sont représentés sous la forme de poids synaptiques $w_{ij}^{(1)}$ et $w_{jk}^{(2)}$. Les biais dans la première couche sont représentés par des poids $b_i^{(1)}$ qui relient des entrées supplémentaires avec une activation égale à un (1). De même, les biais dans la couche de sortie sont représentés comme des poids $b_j^{(2)}$ d'un neurone d'entrée supplémentaire, qui connectent une activation également fixée à un (1). On peut donc écrire l'expression analytique pour les éléments générés par le perceptron monocouche.

L'entrée d'un neurone dans la première couche est obtenue d'abord en formant une combinaison linéaire des signaux de l'ensemble d'entrée, tout en ajoutant le biais pour donner une réponse définie par :

$$y_j^{(1)} = \phi^{(1)}(v_j^{(1)}) = \phi^{(1)}\left(\sum_{i=0}^n \alpha_{ji}^{(1)} x_i\right) \quad j = 1, \dots, h_1, \quad (1.11)$$

avec

$$v_j^{(1)} = b_j^{(1)} + \sum_{i=1}^n w_{ji}^{(1)} x_i = \sum_{l=0}^n \alpha_{jl}^{(1)} x_l \quad j = 1, \dots, h_1, \quad (1.12)$$

$\alpha_{ji}^{(1)}$ pour $i = 0, \dots, n$ désigne le poids synaptique du neurone j dans la couche (1) qui provient de l'entrée i . Les sorties du réseau sont obtenues en transformant les signaux de sortie des neurones dans la première couche par les neurones dans la couche de sortie. Ainsi, l'entrée pour chaque neurone de sortie k est obtenue en formant une combinaison linéaire des signaux de sortie des neurones de la première couche :

$$v_k^{(2)} = b_k^{(2)} + \sum_{j=1}^{h_1} w_{kj}^{(2)} y_j^{(1)} = \sum_{j=0}^{h_1} \alpha_{kj}^{(2)} y_j^{(1)} \quad k = 1, \dots, m, \quad (1.13)$$

$\alpha_{kj}^{(2)}$, $j = 0, \dots, h_1$ désigne les poids synaptiques du neurone k dans la couche (2) qui vient de l'entrée j . L'activation du neurone de sortie k est obtenue en transformant la combinaison linéaire dans l'équation (1.13) avec une fonction d'activation $\phi^{(2)}$ pour donner :

$$y_k = y_k^{(2)} = \phi^{(2)}\left(v_k^{(2)}\right) \quad k = 1, \dots, m, \quad (1.14)$$

Notons que la fonction d'activation dans la couche de sortie $\phi^{(2)}$ n'a pas besoin d'être identique à la fonction de la couche cachée $\phi^{(1)}$. Si nous combinons les équations (1.12), (1.13) et (1.14), on aura une expression explicite pour les sorties représentées dans la figure (1.13)

$$y_k\left(x_i, \alpha_{ji}^{(1)}, \alpha_{kj}^{(2)}\right) = \phi^{(2)}\left(\sum_{j=0}^{h_1} \alpha_{kj}^{(2)} y_j^{(1)}\right) = \phi^{(2)}\left(\sum_{j=0}^{h_1} \alpha_{kj}^{(2)} \phi^{(1)}\left(\sum_{i=0}^n \alpha_{ji}^{(1)} x_i\right)\right) \quad k = 1, \dots, m, \quad (1.15)$$

Le Perceptron simple ne peut résoudre que des problèmes où les exemples peuvent être séparés par un hyperplan. Ces problèmes sont dits linéairement séparables [2].

Les perceptrons multicouches (PMC)

Le perceptron multicouche [19] est un réseau de neurones artificiels organisé en couches où l'information voyage dans un seul sens, de la couche d'entrée vers la couche de sortie. De plus, aucun neurone n'est relié à un neurone situé sur la même

couche, en effet, les poids reliant les neurones ne "sautent" pas de couches ainsi, il n'existe pas de poids reliant un neurone de la couche l à un neurone de la couche $(l + 2)$; $(l + 3)$. Dans le cas général, un perceptron multicouche peut posséder un nombre de couches quelconque et un nombre de neurones (ou d'entrées) par couche également quelconque. Par conséquent, le réseau est regroupé en une séquences qui contient $(c + 1)$ couches $(L^1), \dots, (L^i), \dots, (L^{(c+1)})$ de sorte que les neurones dans n'importe quelle couche (L^i) sont reliés seulement aux neurones dans la couche qui suit par des connexions pondérées. Ce sont les poids de ces connexions qui gouvernent le fonctionnement du réseau et programment une application de l'espace des entrées vers l'espace des sorties à l'aide d'une transformation non linéaire. La création d'un perceptron multicouche pour résoudre un problème donné passe donc par l'inférence de la meilleure application possible telle que définie par un ensemble de données d'apprentissage constituées de paires de vecteurs d'entrées et de sorties désirées. Cette inférence peut se faire, entre autre, par l'algorithme dit de rétropropagation [20].

La figure (1.14) donne l'exemple d'un perceptron multicouches comportant n entrées, h_1 neurones sur la première couche et m neurones sur la couche de sortie. Les sorties des neurones de la dernière couche correspondent toujours aux sorties du système.

Le perceptron multicouche est un des réseaux de neurones les plus utilisés pour des problèmes d'approximation, de classification et de prédiction. Il est habituellement constitué de deux ou trois couches de neurones totalement connectés. Ce type de réseau est dans la famille générale des réseaux à propagation vers l'avant c'est à dire qu'en mode normal d'utilisation, l'information se propage dans un sens unique, des entrées vers les sorties.

La détermination de l'architecture du réseau de neurone artificiel est une étape cruciale. C'est à partir de l'architecture du réseau qu'est introduite l'influence de chaque paramètre dans le système ayant pour but de produire des modélisations fiables. Le nombre de couches, le nombre de neurones par couche et les différentes

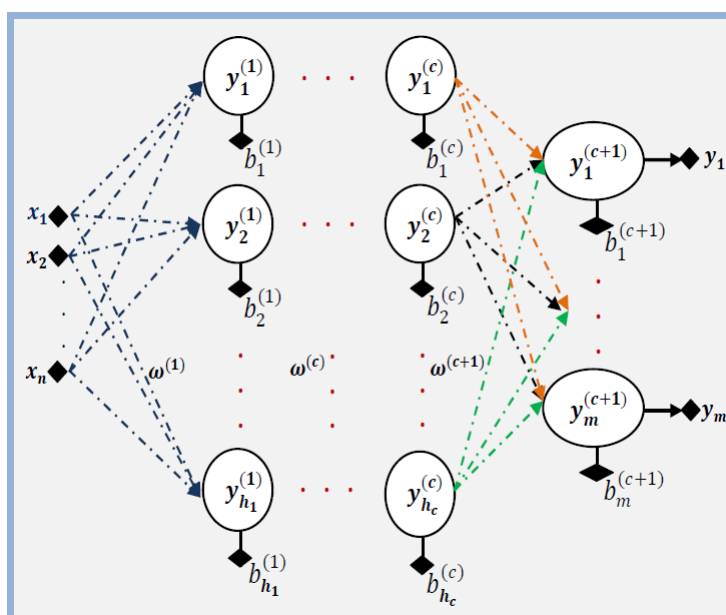


Figure 1.14 – Représentation d’une architecture d’un réseau multicouches de type feedforward [2].

fonctions d’activation déterminent l’architecture du réseau de neurone artificiel [21]. Tous ces paramètres sont déterminés selon le problème traité. Le vecteur $x(x_1, x_2, \dots, x_n)$ présente les entrées de n neurones d’entrées. Il est fixé selon le cas d’étude. Le vecteur $y(y_1, y_2, \dots, y_m)$ présente les sorties calculés de m neurones par le réseau de neurone artificiel dans la couche de sortie. Le nombre de neurones dans les couches cachées est déterminé selon l’apprentissage de notre réseau en évitant le sur-apprentissage qui peut être causé par un trop grand nombre de neurones cachés.

La première couche reçoit les variables d’entrées (x_i) à travers les neurones d’entrées i , les transforme avec la fonction d’activation ϕ sur le neurone d’entrée et les envoie aux neurones j de la première couche cachée. La deuxième couche cachée se compose des neurones de traitement qui reçoivent les sommes pondérées issues de la couche précédente, puis effectuent leurs transformations (avec la fonction d’activation) et finalement les transfèrent à la couche suivante (cachée ou sortie selon l’architecture du réseau de neurone déterminée).

$$v_{ij} = \sum_i^n x_i \cdot w_{ij} + b_j, \quad (1.16)$$

Avec : n est le nombre d'éléments d'entrée ; x_i est la valeur la sortie du i^{eme} neurone de la couche précédente ; w_{ij} est la valeur du poids entre le neurone i et le neurone j ; b_j est le biais.

Les PMC fonctionnent le plus souvent avec une seule couche cachée qui offre déjà un degré de liberté suffisant et se rapproche d'une fonction de type non-linéaire [22, 23]. Quel que soit le nombre de couches cachées dans le PMC, l'équation de production dans un PMC à trois couches (n neurones dans la couche d'entrée, m neurones dans la couche cachée et p neurones dans la couche de sortie) est la suivante :

$$y_k = \phi_k \left(\sum_{j=1}^m w_{jk} \cdot \phi_j \left(\sum_{i=1}^n x_i \cdot w_{ij} + b_j \right) \right) + b_k, \quad (1.17)$$

Avec y_k est l'une des sorties du système ; ϕ_k et ϕ_j sont successivement les fonctions d'activation du neurone k de la couche de sortie et du neurone j de la couche cachée ; w_{jk} et w_{ij} sont successivement les poids entre le j^{eme} neurone et le k^{eme} neurone de sortie et le poids entre le i^{eme} neurone et le j^{eme} neurone.

Le PMC composé d'une couche cachée, avec des fonctions d'activation sigmoïde et une couche de sortie constituée d'un neurone linéaire possède la propriété d'approximateur universel parcimonieux.

Les réseaux de neurones récurrents(RNN)

Les réseaux de neurones récurrents (ou RNN pour Recurrent Neural Networks) [24] sont une extension du PMC dans lesquels des connexions latérales relient les neurones d'une même couche intermédiaire et possédant ainsi des connexions récurrentes au sens où elles conservent des informations en mémoire, le réseau multicouche récurrent est particulièrement adopté pour estimer des processus et systèmes dynamiques non linéaires car : il présente de manière inhérente un caractère dynamique, chaque neurone introduit une portion de non linéarité, il est capable

d'apprendre et peut donc s'adapter en ligne à toute variation de paramètres du système ou tout changement de son environnement, En effet, ils peuvent prendre en compte à un instant t un certain nombre d'états passés. Pour cette raison, les RNNs sont particulièrement adaptés aux applications faisant intervenir le contexte, et plus particulièrement au traitement des séquences temporelles comme l'apprentissage et la génération de signaux, c'est à dire quand les données forment une suite et ne sont pas indépendantes les unes des autres la reconnaissance de la voix par les machines (Speech Recognition) ou la compréhension et la génération de texte (Natural Langage Processing)

Pour schématiser, comme on peut le voir dans la figure (1.15), les neurones des couches intermédiaires reçoivent en entrée les signaux provenant des neurones de la couche précédente combinés à leurs propres signaux de sortie, retardés d'un pas de temps δ . Les connexions réentrantes introduisent donc une notion de mémoire à court terme [25]. Ces réseaux montrent qu'il est possible d'augmenter significativement les performances du PMC traditionnel en modifiant légèrement son principe de fonctionnement.

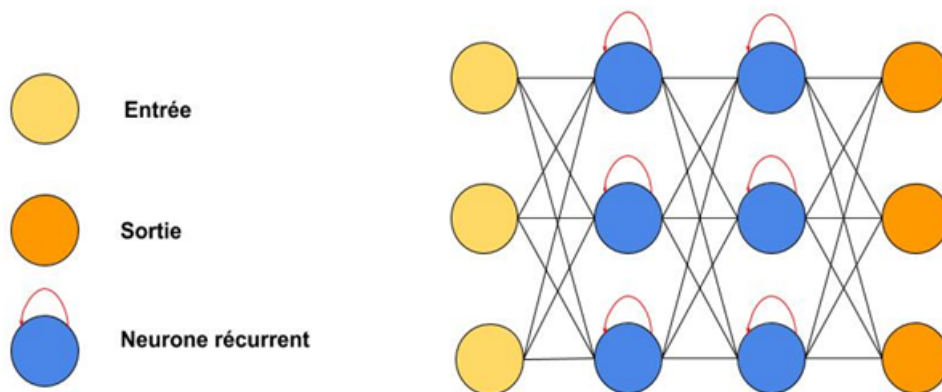


Figure 1.15 – Représentation d'un réseau récurrent.

Les réseaux RBF

Les réseaux RBF [26, 27] (pour Radial Basis Functions) sont une autre alternative au PMC. Ils reposent sur le fait que toute fonction continue peut-être approchée par une somme ou combinaison linéaire de fonctions gaussiennes.

Un réseau RBF est un réseau en trois couches : la couche d'entrée et la couche

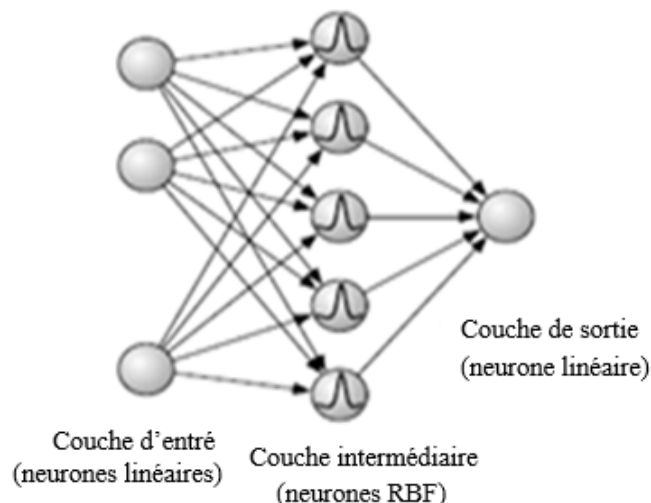


Figure 1.16 – Représentation d'un réseau RBF : perceptron multicouche amélioré dans lequel les neurones de la couche intermédiaire sont des neurones à base radiale.

de sortie contiennent des neurones à fonction linéaire. La couche intermédiaire contient des neurones à fonction gaussienne. Pour une entrée donnée, chacun de ces neurones répond en sortie par la valeur de la gaussienne en ce point. La sortie du réseau correspond à la combinaison linéaire de ces gaussiennes. Leur principal intérêt est de simplifier considérablement l'apprentissage en appliquant de manière locale une régression sur les paramètres de chaque gaussienne ce qui les rend les plus simples à utiliser et les plus rapides.

1.8.2 Les réseaux à interconnexion complète

Les réseaux à interconnexion complète sont une autre possibilité d'organisation, dans laquelle chaque couple de neurones est relié par une connexion. On ne dis-

tingue plus de couche d'entrée ou de couche de sortie car, l'enjeu n'est plus la propagation avant du signal. Ces neurones possèdent généralement une couche principale dans laquelle l'intérêt est d'associer les neurones entre eux. On parle d'ailleurs communément de réseaux associatifs. Ils servent à fabriquer des mémoires associatives, pour des tâches de reconnaissance des formes, une mémoire associative est une mémoire adressable qui permet de retrouver une information stockée à partir d'informations qui lui sont associées [28]. Fondamentalement, le réseau associatif a donc pour but de créer et stocker des motifs de connaissances pour ensuite être capable de les retrouver à partir d'indices. Les motifs de connaissances sont formés par la création de connexions entre les neurones.

Le réseau de Hopfield

Le réseau Hopfield [6] est un réseau neuronal artificiel à une seule couche qui peut être utilisé pour rappeler des modèles qui y ont été stockés. Le réseau Hopfield peut servir de mémoire associative adressable par le contenu car, lorsqu'un motif d'entrée bruyant lui est attribué, il converge vers l'un des motifs avec lesquels il a été entraîné, qui correspond le mieux au motif d'entrée. Chaque neurone est connecté à tous les autres neurones et une valeur est attribuée à cette connexion appelée poids de la connexion. Un moyen simple de former un réseau Hopfield consiste à utiliser la règle d'apprentissage Hebbienne, il n'y a pas de distinction entre neurone d'entrée et neurone de sortie ; seule importe la matrice des poids les neurones sont en effet reliés deux à deux par des connexions symétriques. Une chose importante à prendre en compte est le nombre de modèles qu'un réseau peut stocker. Ce nombre est environ 0.14 fois le nombre de neurones du réseau. Par exemple, un réseau de 100 neurones peut stocker 14 modèles.

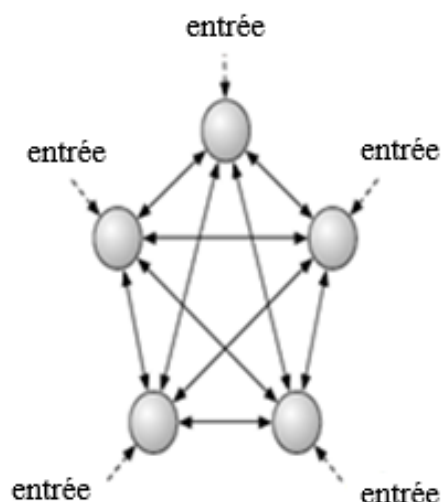


Figure 1.17 – Représentation du réseau de Hopfield.

Le réseau de Kohonen

Les réseaux de Kohonen [16] sont des cartes auto-organisatrices (SOM) avec une variation dans la technique d'apprentissage, ils permettent une classification raffinée. La figure (1.18) présente l'architecture classique d'un réseau de Kohonen. Celui-ci se compose obligatoirement de deux couches : une couche dite des caractéristiques (ou feature map) qui contient un ensemble de neurones totalement interconnectés, et une couche d'entrée qui contient des neurones d'entrée.

De même que pour les autres réseaux associatifs, les réseaux de Kohonen ont été conçus pour faire de la classification. Leur but est de séparer un certain nombre de classes durant l'apprentissage, afin d'être en mesure de discriminer chaque nouvelle entrée lors de l'exploitation du réseau. L'apprentissage est non supervisé, ce qui signifie qu'il consiste uniquement à fournir aux neurones d'entrée des vecteurs. Ensuite, un processus détermine automatiquement si chacun des vecteurs appartient à une classe donnée ou s'il s'agit d'une nouvelle classe à créer. A l'issue de l'apprentissage, le réseau est capable de trier chaque nouvelle entrée dans la classe la plus appropriée. Cet apprentissage, non supervisé et basé sur la

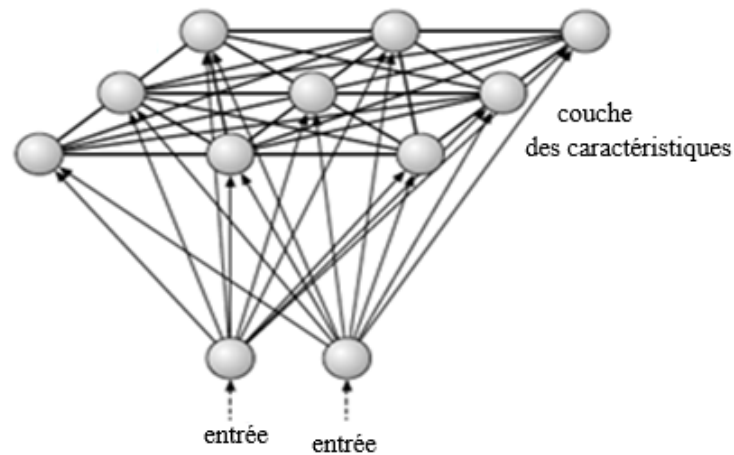


Figure 1.18 – Le réseau de Kohonen.

compétition, est à l'origine des termes sous lesquels on retrouve usuellement les réseaux de Kohonen : carte auto-organisatrice ou auto organisationnelle, ou bien encore SOM (Self Organizing Map).

Les ART (Adaptive Resonance Theory)

Comment pouvons-nous continuer à apprendre de nouvelles informations tout en assurant la relative stabilité des connaissances antérieures ? C'est dans l'optique de répondre à cette question que Carpenter et Grossberg [29] ont élaboré la théorie de la résonance adaptative qu'ils formaliseront à travers leur réseaux de neurones. La figure (1.19) présente le modèle général de ces réseaux.

Le réseau ART est composé de deux niveaux F_1 et F_2 . Le niveau F_1 accepte les vecteurs d'entrée. Le niveau F_2 contient les classes ou catégories que le réseau a créé ; chaque neurone correspond à une classe distincte des autres.

L'apprentissage par résonance utilise la règle de Grossberg pour renforcer les connexions entre un ensemble de caractéristiques du signal sur la couche F_1 et une classe de la couche F_2 . Le réseau choisit la classe la plus appropriée en regard de l'apprentissage déjà effectué, il est capable de s'adapter à des entrées non familières

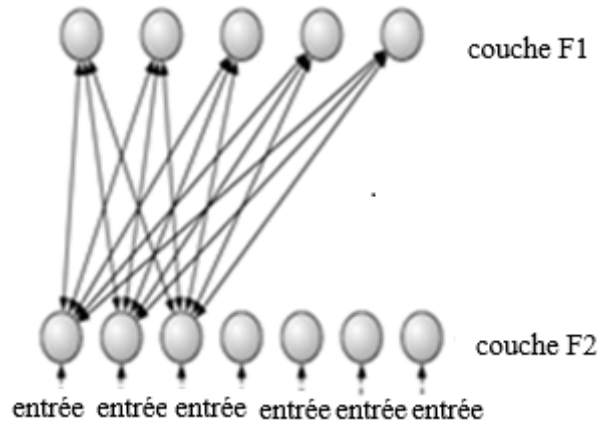


Figure 1.19 – Le réseau ART de Grossberg et Carpenter.

en construisant de nouvelles catégories ou classes mais aussi d'adapter les classes déjà apprises tout en dégradant peu les informations mémorisées

1.9 Conclusion

Dans ce chapitre, nous avons introduit les définitions essentielles relatives aux réseaux de neurones formels. Bien que les réseaux de neurones tirent leur origine de la neurobiologie, ils doivent être considérés comme des outils mathématiques et statistiques sous forme de fonctions paramétrées ou le calcul est fondé sur la propagation de l'information entre des unités de calcul élémentaires appelés neurones formels interconnectés par des liaisons pondérées. Les chemins de calcul entre ces liaisons sont très courts et les neurones peuvent être vus comme des éléments de décision régénérant en permanence le signal. Leur interconnexion permet d'effectuer un calcul global complexe bien adapté qui permet de réaliser, de manière simple et efficace, des modèles statiques ou dynamiques bien précis. Le champ d'application est donc très vaste et regroupe la classification de patterns, la reconnaissance de la parole, la vision des machines, la robotique, le traitement de signal, les problèmes d'optimisation et de calcul.

2 Chapitre 2 : Méthodologie des réseaux multicouches et algorithmes d'optimisation

2.1 Introduction

Les réseaux de neurones artificiels tirent leur origine des neurones biologiques qui composent le cerveau humain. Leur conception vise à reproduire une de ses caractéristiques les plus importantes : la capacité d'apprendre. Chaque neurone artificiel peut fonctionner indépendamment sans se préoccuper du fonctionnement global du réseau. Néanmoins, malgré cette façon décentralisée de traiter l'information, le réseau neuronal dans sa totalité demeure capable de calculer des gradients ce qui revient à déterminer comment ajuster au mieux ses paramètres. L'estimation de ces paramètres est effectuée au cours d'une étape de développement du réseau appelée : processus d'apprentissage. Lorsque nous intégrons un exemple d'apprentissage au réseau, les neurones passent à un état actif qui entraîne un changement itératif dans ses termes synaptiques. Ce processus permet au réseau de s'auto organiser sur la base des exemples présentés comme unités de vecteurs d'entrée et de sortie. L'apprentissage et l'optimisation des réseaux de neurones [30] sont des étapes décisives à accomplir. Pour cette raison, des variétés d'algorithmes d'apprentissage ont été proposées, nous indiquons pour chacun d'entre eux, leur intérêt pratique et leurs limites. Les différents exemples présentés dans ce chapitre montrent que l'efficacité de l'apprentissage augmente avec le nombre de points d'apprentissage. Ceci est un facteur très important qu'il faut souligner : quelque soit l'algorithme choisit, la qualité de l'apprenissage des réseaux est d'autant meilleure que l'on dispose d'un ensemble d'apprentissage riche en exemples [31]

2.2 Méthodologie des réseaux multicouches et algorithmes d'optimisation

A partir d'une architecture donnée de réseau de neurones , il est possible d'engendrer une famille de fonctions paramétrées par les valeurs des coefficients

du réseau (ou poids synaptiques) [32]. L'objectif de la phase d'apprentissage des réseaux de neurones est de trouver, parmi toutes ces fonctions, celle qui s'approche le plus possible de la régression (fonction génératrice des exemples). Pour cela, il faut définir une fonction de cout qui mesure l'écart entre la sortie du modèle (fonction réalisée par le réseau de neurones) et la sortie désirée constitué par les poids du réseau. La fonction de cout est une fonction scalaire qui dépend du vecteur de paramètres noté α du modèle. Plus la valeur de la fonction de cout est petite, plus le modèle reproduit fidèlement les observations utilisées pour l'apprentissage. Les différents algorithmes d'apprentissage cherchent donc à trouver le point, dans l'espace des paramètres, pour lequel la fonction de cout est minimale.

Pour un exemple i d'un ensemble d'observations E , la fonction de cout des moindres carrés est égale à la somme, sur les N_s neurones de la couche de sortie, des carrés des écarts entre la sortie du modèle (sortie du réseau de neurone $:y_i$) et la sortie désirée (grandeur mesurée notée d_i) [33]. comme la sortie du réseau de neurone dépend du vecteur paramètre α , la fonction de cout en dépend également. On la note $\varepsilon^i(\alpha)$:

$$\varepsilon^i(\alpha) = \frac{1}{2} \sum_{q=1}^{N_s} (d_q^i - y_q^i(\alpha))^2. \quad (2.1)$$

A partir d'une fonction de cout dépendant du vecteur de paramètres α et des exemples de l'ensemble d'apprentissage, il faut choisir l'algorithme d'optimisation qui permettra d'estimer le vecteur des paramètres pour lequel la fonction de cout choisie est minimale. De nombreux algorithmes ont été proposés ; dans ce paragraphe, nous décrivons les différents algorithmes utilisés pour l'apprentissage des réseaux de neurones.

2.3 Algorithme d'apprentissage

La procédure utilisée pour mener à bien le processus d'entraînement est appelée: algorithme d'apprentissage. Ce dernier est appliqué au réseau afin d'obtenir

une performance souhaitée. Le type d'apprentissage est déterminé par la manière dont l'ajustement des paramètres dans le réseau neuronal a lieu [34]. Pour sa réalisation il faut définir :

1. Un ensemble d'exemples d'apprentissage : les réseaux de neurones sont des fonctions paramétrées, utilisées pour réaliser des modèles statistiques à partir d'exemples (dans le cas de la classification) ou de mesures (dans le cas de la modélisation); leurs paramètres sont calculés à partir de ces exemples ou couples (entrée, sortie).
2. La définition d'une fonction de cout qui mesure l'écart entre les sorties du réseau de neurone et les sorties désirées (dans le cas de la classification) ou les valeurs mesurées (dans le cas de la modélisation) présentes dans l'ensemble d'apprentissage.
3. Un algorithme de minimisation de la fonction de cout par rapport aux paramètres.

Le développement des techniques d'accélération de convergence des algorithmes d'apprentissage a mené à deux types différents d'algorithmes : les algorithmes de premier et second ordres.

2.3.1 Méthodes du premier ordre

a. Évaluation du gradient de la fonction quadratique moyenne par descente de gradient

Les algorithmes d'apprentissage fondés sur l'évaluation du gradient de la fonction de coût $\varepsilon(\alpha)$ par rapport aux paramètres procèdent à la minimisation de manière itérative. $\varepsilon(\alpha)$ est une fonction scalaire à variable vectorielle (le vecteur α des paramètres à ajuster). Son gradient est donc un vecteur défini par [35] :

$$\nabla \varepsilon = \left(\frac{\partial \varepsilon}{\partial \alpha_1}, \frac{\partial \varepsilon}{\partial \alpha_2}, \dots, \frac{\partial \varepsilon}{\partial \alpha_m} \right)^\tau, \quad (2.2)$$

où m est le nombre de paramètres inconnus.

Le principe des algorithmes de gradient repose sur le fait qu'un minimum de la fonction de coût est atteint si sa dérivée (son gradient) est nul. Leur déroulement suit le schéma suivant :

initialisation

- **A l'itération 0** : Initialiser le vecteur des paramètres à α_0 . Cette initialisation de α peut avoir une grande influence sur l'issue de l'apprentissage
- Choisir le pas μ_k
- Fixer un seuil de convergence $\delta\varepsilon > 0$

Itérations (boucle sur k)

- **A la k^{eme} itération** : calculer la fonction de coût $\varepsilon(\alpha^k)$ et la norme du gradient $\nabla\varepsilon(\alpha^k)$.
- Choisir comme direction de descente $d^k = -\nabla\varepsilon(\alpha^k)$
- Déterminer α^{k+1} selon la formule

$$\alpha^{k+1} = \alpha^k - \mu_k \nabla\varepsilon(\alpha^k), \quad (2.3)$$

- Test de convergence

$$\frac{|\varepsilon(\alpha^{k+1}) - \varepsilon(\alpha^k)|}{\varepsilon(\alpha^0)} \leq \delta. \quad (2.4)$$

En pratique, la méthode du gradient simple peut être efficace lorsque l'on est loin du minimum de ε . Quand on s'en approche, la norme du gradient diminue et donc l'algorithme progresse plus lentement. A ce moment, on peut utiliser une méthode de gradient plus efficace. Un (réglage) du pas de gradient μ_k est nécessaire : en effet, une petite valeur de ce paramètre ralentit la progression de l'algorithme ; en revanche une grande valeur aboutit généralement à un phénomène d'oscillation autour de la solution.

Diverses heuristiques, plus ou moins efficaces, ont été proposées.

Techniques de réglage du pas

- **Technique du pas constant** elle consiste à adopter un pas constant $\mu_k = \mu$ tout au long de l'algorithme. Elle est très simple mais peu efficace puisqu'elle ne prend pas en considération la décroissance de la norme du gradient.
- **Technique du pas asservi** on peut asservir le pas à l'aide de la norme du gradient de sorte que le pas évolue en sens inverse de celle-ci. A chaque étape, le pas peut être calculé par :

$$\mu_k = \frac{\mu}{1 + \|\nabla \varepsilon\|}. \quad (2.5)$$

où μ est un paramètre constant. Lors de l'utilisation de cette technique, nous avons adopté la valeur $\mu = 10^{-3}$ qui s'est révélée très souvent satisfaisante. Le numérateur est augmenté du nombre 1 afin d'éviter une instabilité numérique au moment de la division dans le cas où la norme du gradient devient très proche du zéro. Cette technique offre un bon compromis du point de vue de la simplicité et de l'efficacité.

b. Évaluation du gradient de la fonction quadratique moyenne par rétro-propagation

La règle de la rétropropagation du gradient de l'erreur [36] consiste à propager l'erreur obtenue à une unité de sortie d'un réseau à couches comportant une ou plusieurs couches cachées à travers le réseau par descente du gradient dans le sens inverse de la propagation des activations. Cet algorithme permet de réaliser un apprentissage du réseau de neurones. On cherche à obtenir du réseau une réponse préétablie comme étant correcte. On dispose d'une base de connaissance de type entrée (x), sortie attendue (d). On compare ensuite la sortie (y) à la sortie attendue. On introduit une fonction $e = (d-y)$ qui est la fonction d'erreur que l'on va chercher à minimiser en modifiant les poids du réseau.

Considérons un réseau PMC comportant plusieurs couches cachées et une couche de sortie à plusieurs neurones. On utilise les notations suivantes :

m : Nombre de neurones sur la couche k .

p : Nombre d'entrées de la couche k .

d : Sortie attendue à la dernière couche du réseau.

y_k : Sortie de la couche k du réseau.

w : Matrice des poids de la couche k .

b : Vecteur biais de la couche k .

ϕ : Fonction d'activation de la couche k .

En général tous les neurones cachés ont la même fonction d'activation qui est souvent une fonction sigmoïde. Les neurones de sortie ont également la même fonction d'activation, en général linéaire (par exemple $f(\nu) = \nu$).

$$y_j^{(k)} = \phi(v_j^{(k)}(w, b)), \quad (2.6)$$

$$v_j^{(k)} = \sum_{i=1}^p w_{ij}^{(k)} y_i^{(k-1)} - b_j^{(k)}, \quad (2.7)$$

L'erreur entre la sortie désirée d_i et la sortie y_i produite par le neurone de sortie d'indice i s'exprime par :

$$e_i = d_i - y_i, \quad (2.8)$$

L'erreur quadratique moyenne du réseau est :

$$\varepsilon = \frac{1}{2} \sum_{i \in S} [e_i]^2. \quad (2.9)$$

Où S est l'ensemble des indices des neurones de sortie.

L'objectif de l'apprentissage est de minimiser le critère ε défini par la relation (2.9) dans laquelle α représente le vecteur des poids du réseau d'éléments $\alpha_{i,j}$. L'algorithme de rétropropagation du gradient est un algorithme itératif permettant

de modifier les paramètres.

A fin de minimiser l'erreur quadratique moyenne du réseau sur l'ensemble des données d'apprentissage, on suit une technique de descente du gradient :

$$\alpha_{ij}^{k+1} = \alpha_{ij}^k - \mu \frac{\partial \varepsilon}{\partial \alpha_{ij}^k}, \quad (2.10)$$

$$\Delta w_{i,j} = -\mu \frac{\partial \varepsilon}{\partial w_{i,j}}(w, b), \quad (2.11)$$

$$\frac{\partial \varepsilon}{\partial w_{i,j}}(w, b) = \frac{\partial \varepsilon}{\partial v_j^k} \frac{\partial v_j^k}{\partial w_{i,j}}, \quad (2.12)$$

On va calculer séparément les deux facteurs:

$$\frac{\partial v_j^k}{\partial w_{i,j}} = \frac{\partial}{\partial w_{i,j}} [w y^{k-1} - b^k] = y_j^{k-1}, \quad (2.13)$$

On note

$$\delta_j^k = \frac{\partial \varepsilon}{\partial v_j^k}, \quad (2.14)$$

La quantité δ est la rétropropagation des erreurs en arrière et est considérée pour chaque neurone dans le réseau. En remplaçant (2.14) dans (2.12), on obtient alors

$$\Delta w_{i,j} = -\mu \delta_j^k y_j^{k-1}, \quad (2.15)$$

En refaisant exactement les mêmes calculs en dérivant par rapport à b_j (en s'intéressant donc à la modification du biais), on obtient

$$\Delta b_j = -\mu \delta_j^k, \quad (2.16)$$

Calcul de la sensibilité : La sensibilité de la dernière couche est exprimée par

$$\delta_i^N = \frac{\partial \varepsilon}{\partial v_i^N} = \frac{\partial}{\partial v_i^N} [d_i - \phi(v_i^N(w, b))]^2 = -2(d_i - y_i) \phi'(v_i^N), \quad (2.17)$$

Et comme

$$v^{K+1} = w^{k+1} \phi^k(v^k) - b^{k+1}, \quad (2.18)$$

$$\delta_i^k = \frac{\partial \varepsilon}{\partial v_i^k} = \frac{\partial \varepsilon}{\partial v_j^{k+1}} \frac{\partial v_j^{k+1}}{\partial v_i^k} = \delta_j^{k+1} \frac{\partial v_j^{k+1}}{\partial v_i^k} \quad (2.19)$$

Reste à calculer le second facteur de cette expression

$$\frac{\partial v_j^{k+1}}{\partial v_i^k} = \frac{\partial}{\partial v_i^k} \left[\sum_l w_{j,l}^{k+1} \phi^k(v_l^k) - b_j^{k+1} \right] = w_{j,i}^{k+1} \phi'_k(v_i^k) \quad (2.20)$$

On a donc une relation de récurrence entre δ^k et δ^{k+1} (d'où le terme "rétropropagation")

$$\delta^K = \phi'_k(v_i^k W^{k+1}) \delta^{K+1} \quad (2.21)$$

$$\delta^N = -2\phi'_N(v^N) (d - y) \quad (2.22)$$

Connaissant les gradients locaux des neurones de la dernière couche cachée, on peut calculer pas à pas les gradients locaux des neurones des couches cachées précédentes en utilisant de façon récursive la relation (2.21) L'algorithme de rétropropagation du gradient peut ainsi se résumer en deux étapes :

1. Pendant la première étape, l'information (les entrées) transite de l'entrée du réseau vers sa sortie, les poids des connexions étant maintenus fixes. On calcule alors les erreurs à la sortie du réseau.
2. Pendant la deuxième étape l'information (les erreurs à la sortie du réseau) transite de la sortie vers l'entrée et permet la modification des poids par la relation (2.15) où les gradients locaux δ^N sont calculés par la relation (2.22) pour les neurones de sortie et (2.21) pour les neurones cachés.

Un "réglage" du pas de gradient δ est nécessaire: en effet, une petite valeur de ce paramètre ralentit la progression de l'algorithme; en revanche une grande valeur aboutit généralement à un phénomène d'oscillation autour de la solution. Cette méthode est très utilisée. Elle a pour avantages une grande facilité de réalisation et

une grande robustesse. La méthode est efficace loin d'un minimum mais la vitesse de convergence diminue lorsque l'on s'approche du minimum (la modification des paramètres est proportionnelle au gradient de la fonction de cout qui tend vers 0).

2.3.2 Méthodes du second ordre

Les algorithmes du second ordre convergent plus rapidement par rapport aux algorithmes de premier ordre [37]. Ils estiment automatiquement le taux d'apprentissage dans chaque direction du gradient en utilisant la matrice Hessienne (2.24). En combinant la vitesse d'apprentissage des algorithmes du second ordre et la stabilité de l'algorithme de rétropropagation on obtient des variétés d'algorithmes d'apprentissages proposés pour ajuster les paramètres pondérés des réseaux de neurones artificiels [38, 39]. Ces algorithmes sont basés sur le développement limité au second ordre au voisinage du vecteur α^k obtenu à l'itération k de la fonction de cout $\varepsilon(\alpha)$:

$$\varepsilon(\alpha^{k+1}) \approx \varepsilon(\alpha^k) + \nabla \varepsilon(\alpha^k) (\alpha^{k+1} - \alpha^k) + \frac{1}{2} (\alpha^{k+1} - \alpha^k) H(\alpha^k) (\alpha^{k+1} - \alpha^k) \quad (2.23)$$

Où H est la matrice hessienne de la fonction de cout ε , définie par

$$H = \begin{bmatrix} \frac{\partial^2 \varepsilon}{(\partial \alpha_{11}^{(k)})^2} & \frac{\partial^2 \varepsilon}{\partial \alpha_{11}^{(k)} \partial \alpha_{12}^{(k)}} & \cdots & \frac{\partial^2 \varepsilon}{\partial \alpha_{11}^{(k)} \partial \alpha_{1n}^{(k)}} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 \varepsilon}{\partial \alpha_{mm}^{(k)} \partial \alpha_{m1}^{(k)}} & \frac{\partial^2 \varepsilon}{\partial \alpha_{mm}^{(k)} \partial \alpha_{m2}^{(k)}} & \cdots & \frac{\partial^2 \varepsilon}{(\partial \alpha_{mn}^{(k)})^2} \end{bmatrix} \quad (2.24)$$

a. Méthode de Newton

La méthode de Newton [40] utilise la courbure (dérivée seconde) de la fonction de cout pour atteindre plus rapidement le minimum. Supposons que α^{k+1} soit un

minimum de la fonction de cout. Le gradient de celle-ci est donc nul, ce qui s'écrit

$$0 = \nabla \varepsilon (\alpha^k) + H (\alpha^k) (\alpha^{k+1} - \alpha^k) \quad (2.25)$$

On obtient donc l'estimation itérative recherchée

$$\alpha^{k+1} = \alpha^k - H^{-1} (\alpha^k) \nabla \varepsilon (\alpha^k) \quad (2.26)$$

Si la matrice hessienne est inversible

La direction de descente est dans ce cas

$$d_{k+1} = -H^{-1} (\alpha^k) \nabla \varepsilon (\alpha^k) \quad (2.27)$$

Pour que la direction de mise à jour des paramètres corresponde effectivement à une direction de descente, il est nécessaire que la matrice Hessienne soit définie positive (ce qui assure par ailleurs son inversibilité, et le fait que l'inverse est également défini positif). Cette méthode de recherche itérative du minimum, présente de très bonnes propriétés de convergence au voisinage du minimum, mais nécessite l'inversion d'une matrice, ce qui peut s'avérer lourd en termes de temps de calcul. Cependant, cette augmentation du temps de calcul pour une itération (par comparaison à la formule de mise à jour des paramètres) est compensée par la diminution de nombre d'itérations nécessaires.

b. Méthode de Levenberg-Marquardt

Cette méthode est très proche de la méthode de Newton décrite précédemment. La seule différence réside dans l'introduction d'un paramètre λ , appelé paramètre de Levenberg-Marquardt, permettant de stabiliser la méthode de Newton. Ce paramètre est actualisé automatiquement en fonction de la convergence de chaque itération [41]. Une stabilisation est possible grâce à un procédé réitératif (si une itération diverge, on la recommence au départ en augmentant le paramètre λ

jusqu'à obtenir une itération convergente). L'algorithme de Levenberg-Marquardt obéit à la formule suivante de mise à jour des paramètres

$$\alpha^{k+1} = \alpha^k - [\mathbf{H}(\alpha^k) + \lambda_{k+1}\mathbf{I}]^{-1} \nabla \varepsilon(\alpha^k) \quad (2.28)$$

Où $\mathbf{H}(\alpha^k)$ est la matrice hessienne de la fonction de cout ε , \mathbf{I} est la matrice identité. La méthode de Levenberg Marquardt est astucieuse car elle s'adapte d'elle-même à la forme de la fonction du cout. Elle choisit automatiquement un compromis entre la direction du gradient et la direction de Newton : pour de petites valeurs du pas, cette méthode s'approche de celle de Newton. Inversement, pour de grandes valeurs du pas λ_{k+1} , l'algorithme Levenberg Marquardt est équivalent à l'application de la règle du gradient simple avec un pas de $\frac{1}{\lambda_{k+1}}$.

Nous choisissons une valeur initiale λ_0 qui est modifiée durant l'optimisation. A chaque itération, on calcule la fonction de cout $\varepsilon(\alpha)$ avec la valeur de λ précédente si la fonction de cout diminue, on effectue la modification des paramètres et on diminue λ . (par exemple : division par 10); si la fonction de cout croit, on cherche à se rapprocher du gradient et on augmente λ (multiplication par 10) jusqu'à ce que le cout diminue. Il suffit de spécifier les critères d'arrêt et l'algorithme adapte λ , on constate qu'en début d'optimisation, λ augmente (la direction de descente est presque celle du gradient) puis diminue au voisinage du minimum (la direction de descente est presque celle de Newton). Cette souplesse se paie par un temps de calcul sensiblement supérieur [42].

c. L'algorithme de BFGS

Cet algorithme (du nom de ses inventeurs Broyden, Fletcher, Glodfarb et Shanno) est fondé sur une approximation de la méthode de Newton exposée précédemment. La règle de mise à jour des paramètres est définie comme suit [43]

$$\alpha^{k+1} = \alpha^k - \mu_{k+1} M_{k+1} \nabla \varepsilon(\alpha^k) \quad (2.29)$$

Où μ_{k+1} est appelé pas de descente, et où M_{k+1} est une approximation, calculée itérativement, de l'inverse de la matrice Hessienne, selon la formule suivante

$$M_{k+1} = M_k + \left(1 + \frac{y_k^T M_k \gamma_k}{\delta_k^T \gamma_k} \right) \frac{\delta_k^T \delta_k}{\delta_k^T \gamma_k} - \frac{\delta_k \gamma_k^T M_k + M_k \gamma_k \delta_k^T}{\delta_k^T \gamma_k}, \quad (2.30)$$

Où

$$\gamma_k = \nabla \varepsilon (\alpha^k) - \nabla \varepsilon (\alpha^{k-1}) \quad \text{et} \quad \delta_k = \alpha^k - \alpha^{k-1}, \quad (2.31)$$

La valeur initiale de la matrice M est généralement la matrice identité (valeur à laquelle M_{k+1} sera également réinitialisée au cours de l'algorithme si elle s'avère ne plus être définie positive). La valeur du pas μ_k peut être choisie comme étant la valeur permettant la diminution la plus importante de la fonction de coût. On procède alors de la façon suivante [44]

1. initialiser μ_k à une petite valeur (typiquement 10^{-3}) ;
2. mettre à jour les paramètres conformément à (2.29), puis vérifier que la condition de descente est bien respectée $\varepsilon(\alpha^{k+1}) < \varepsilon(\alpha^k)$;
3. si tel est le cas, rechercher une valeur du pas plus importante (par multiplication par un facteur supérieur à 1), et vérifier que cette valeur conduit à une fonction de coût moindre que précédemment;
4. si la condition de descente n'est pas vérifiée, multiplier le pas par un facteur inférieur à 1, et tester à nouveau la condition de descente. Cette procédure d'ajustement du pas est arrêtée soit si la décroissance de la fonction de coût est inférieure à un seuil prédéfini, soit si le nombre de tentatives atteint une limite prédéfinie.

L'intérêt de cet algorithme de BFGS réside en ce qu'il permet de s'affranchir du calcul de l'inverse de la matrice Hessienne (qui peut lui-même s'avérer délicat dans certains cas), en estimant itérativement une approximation de cette matrice inverse. Cette méthode quasi-newtonienne n'est efficace qu'à proximité du minimum de la fonction de coût : il convient donc de combiner l'utilisation de

cet algorithme avec, dans un premier temps, une approche du minimum par une méthode du premier ordre.

d. Méthode du gradient conjugué

Dans le cadre de la méthode du gradient conjugué, la recherche de la direction est faite tout au long des directions conjuguées [45]. La technique numérique du gradient conjugué profite implicitement de l'information sur le Hessien (H) qui est la matrice des dérivées secondes partielles pour trouver le minimum d'une fonction d'erreur quadratique. Tous les algorithmes numériques de cette catégorie commencent par une recherche dans la direction de la descente du gradient lors de la première itération. Considérons la fonction d'erreur quadratique $\varepsilon(x)$

$$\varepsilon(x) = \frac{1}{2}x^T H x + b^T x + c \quad (2.32)$$

La matrice Hessienne (définie positive) est égale à :

$$H = \nabla^2 \varepsilon(x) \quad (2.33)$$

Nous définissons les directions d_1, \dots, d_k dites conjuguées par rapport à la matrice H si

$$d_i^T H d_j = 0 \quad 1 \leq i, j \leq k \text{ et } i \neq j \quad (2.34)$$

Il est possible d'atteindre un minimum de $\varepsilon(x)$ en effectuant N recherches linéaires le long des droites orientées dans des directions conjuguées (d_1, d_2, \dots, d_N) . En posant

$$g_k = \nabla F(x)|_{x=x_k} = H x_k - b \quad (2.35)$$

On choisit le pas α_k pour minimiser $\varepsilon(x_k)$ dans la direction de d_k

$$\alpha_k = -\frac{g_k^T d_k}{g_k^T H d_k} \quad (2.36)$$

alors

$$\Delta \mathbf{x}_k = \mathbf{x}_{k+1} - \mathbf{x}_k = \alpha_k \mathbf{d}_k \quad (2.37)$$

Ces procédures changent la direction du gradient à chaque itération de l'algorithme. Ce qui nous permet de faire disparaître la matrice Hessienne de l'équation qui définit la condition des vecteurs conjugués. D'après les conditions d'orthogonalité à la variation de gradient, les vecteurs d_j seront conjugués. Il s'agit à chaque itération k de construire une direction de recherche d_k qui est

$$\mathbf{d}_k = -\mathbf{g}_k + \beta_k \mathbf{d}_{k-1} \quad (2.38)$$

Où β_k est calculé de sorte que

$$d_k^T H d_{k-1} = 0 \Rightarrow (-g_k + \beta_k d_{k-1})^T H d_{k-1} = 0 \quad (2.39)$$

Donc

$$\beta_k = \frac{g_k^T H d_{k-1}}{d_{k-1}^T H d_{k-1}} \quad (2.40)$$

Nous avons présenté dans cette partie les algorithmes du second ordre que nous utiliserons dans ce travail (c'est à dire l'algorithme de BFGS et celui de Levenberg Marquardt). La difficulté essentielle lors de l'application de l'algorithme de BFGS réside dans le choix de la condition de passage du gradient simple à la méthode de BFGS. Ce problème ne se pose pas pour l'algorithme de Levenberg Marquardt, mais le volume de calculs nécessaires à chaque itération de cet algorithme croît rapidement avec le nombre de paramètres.

2.4 Intérêt des réseaux de neurones formels

L'intérêt porté aujourd'hui aux RNF dans plusieurs domaines de recherches (biologie, mathématique, physique, optique, électronique, robotique) se justifie par les quelques propriétés fascinantes qu'ils possèdent, à l'instar du cerveau humain,

le parallélisme, la capacité d'adaptation, la mémoire distribuée et l'émergence de la connaissance.

- Le parallélisme permet une rapidité de calcul. La capacité d'apprentissage et d'adaptation est sûrement une propriété, qui aurait besoin d'être mieux formulée. Il est vrai qu'un RNF n'a pas à être programmé et qu'il est doté de capacités d'apprentissage.
- La mémoire est distribuée, c'est à dire qu'elle émane d'un comportement collectif du réseau. Un élément de mémoire n'est pas physiquement localisable comme dans la mémoire d'un ordinateur conventionnel, mais il est réparti sur tout le réseau. La perte d'un élément ne correspond pas à la perte d'un fait mémorisé et le réseau présente donc une certaine tolérance aux défauts.
- Un autre avantage est de permettre une évolution dynamique du réseau vers un état mémorisé dans le cadre d'un environnement de données floues.

2.5 Propriété fondamentale des réseaux de neurones artificiels

Les réseaux de neurones formels, tels que nous les avons définis précédemment possèdent une propriété remarquable qui est à l'origine de leur intérêt pratique dans des domaines très divers, ce sont des approximateurs universels parcimonieux. Cette expression traduit deux propriétés distinctes : d'une part, les réseaux de neurones sont des approximateurs universels et d'autre part, une approximation à l'aide des réseaux de neurones est parcimonieuse

a- L'universalité

Cette propriété a été démontrée par Cybenko [46] et s'énonce de la façon suivante: Toute fonction bornée suffisamment régulière peut être approchée uniformément, avec une précision arbitraire, dans un domaine fini de l'espace de ses variables, par un réseau de neurones comportant une couche de neu-

rones cachés en nombre fini possédant tous la même fonction d'activation, et un neurone de sortie linéaire. Autrement dit : il existe un réseau de neurones avec un nombre fini de neurones cachés, qui est capable d'approcher toute fonction non linéaire suffisamment régulière dans une région W de l'espace de ses variables avec une précision fixée, cela signifie qu'un réseau de neurones est capable d'imiter n'importe quel processus après ajustement de ses paramètres par un apprentissage

b- La parcimonie

La parcimonie est la propriété mathématique qui donne aux réseaux de neurones tout leur intérêt pratique, elle exprime le fait que le réseau a besoin de moins de paramètres ajustables pour atteindre une précision donnée que les approximateurs traditionnels tels que les polynômes ou les séries de Fourier par exemple, c'est à dire, le nombre de paramètres du réseau varie linéairement avec le nombre d'entrées, tandis qu'il varie exponentiellement avec le nombre de variables de la fonction à approcher pour les approxima-teurs usuels[47]. Les réseaux de neurones sont donc avantageux dès que le procédé que l'on cherche à modéliser à plus de deux entrées. La parcimonie des réseaux de neurones a ainsi une conséquence pratique très importante : pour obtenir une précision de modélisation donnée, un réseau de neurones (qui est un modèle non linéaire par rapport à ses paramètres) nécessite un plus petit nombre de mesures qu'un modèle linéaire par rapport à ses paramètres, donc non parcimonieux. Contrairement à une idée reçue, les réseaux de neurones ne sont donc pas des modèles sur paramétrés, bien au contraire, ils sont parcimonieux en nombre de paramètres.

Qualitativement, la propriété de parcimonie peut se comprendre de la manière suivante : lorsque l'approximation est une combinaison linéaire de fonctions élémentaires fixées (des monômes par exemple, où des gaussiennes à centres et écarts-types fixes), on ne peut ajuster que les coefficients de

la combinaison; en revanche, lorsque l'approximation est une combinaison linéaire de fonctions non linéaires à paramètres ajustables (un Perceptron multicouches par exemple), on ajuste à la fois les coefficients de la combinaison et la forme des fonctions que l'on combine. Ainsi, dans un Perceptron multicouches, les poids de la première couche déterminent la forme de chacune des sigmoïdes réalisées par les neurones cachés, et les poids de la seconde couche déterminent une combinaison linéaire de ces fonctions. On conçoit facilement que cette souplesse supplémentaire, conférée par le fait que l'on ajuste la forme des fonctions que l'on superpose, permet d'utiliser un plus petit nombre de fonctions élémentaires, donc un plus petit nombre de paramètres ajustables [48].

2.6 Les différents domaines d'application des réseaux de neurones

Les applications des réseaux de neurones sont très diverses et se multiplient au fil du temps, parmi ces domaines d'applications privilégiés, nous citerons :

1. L'analyse de la parole : les algorithmes d'analyse de signaux ont été considérablement améliorés par les réseaux de neurones profonds y compris pour séparer des signaux audio mélangés, ce que l'on appelle de la séparation de sources.
2. Les théories du traitement du langage naturel : le traitement informatique du langage naturel a commencé avec des modèles statistiques basés sur des chaînes de Markov et des modèles informatiques des grammaires formelles. Cependant, là encore les réseaux de neurones ont apporté des améliorations considérables pour des applications telles que la traduction automatique ou

la génération de phrases pour répondre à des questions. Ce sont ces réseaux de neurones qui sont actuellement utilisés dans la plupart des applications commerciales.

3. En robotiques pour des jeux de stratégie comme les échecs : les réseaux de neurones sont utilisés avec des algorithmes d'apprentissage par renforcement. L'optimisation du réseau est faite en utilisant l'information donnée par une récompense, qui évalue la performance d'une action.

4. L'application des réseaux de neurones à la physique : en physique, on essaye de prédire l'évolution d'un système en essayant de résoudre des équations fondamentales. Cela peut cependant devenir très compliqué lorsque le système n'est pas simple. La prédiction de phénomènes physiques par réseaux de neurones peut s'assimiler à une forme de physique statistique. On prédit un phénomène à partir d'une base de données d'exemples, en essayant de capturer des formes de régularité de cette évolution en exploitant les capacités de régression non linéaire parcimonieuse des réseaux de neurones pour effectuer ces prédictions. Des résultats de plus en plus précis sont obtenus avec des réseaux de neurones, en physique quantique, en dynamique des fluides et science des matériaux.

5. Les réseaux de neurones profonds ont aussi permis de créer de nouveaux modèles stochastiques afin de générer des données. Cela concerne des phénomènes physiques complexes comme la turbulence en dynamique des fluides ou des textures d'images ou d'audio, aussi bien que des données très structurées comme des images de visages ou des musiques.

2.7 Conclusion

Dans ce chapitre, nous avons présenté les principes de la modélisation, les étapes de la conception ainsi que les fonctions paramétrées utilisables et les algorithmes qu'il convient de mettre en oeuvre pour l'ajustement des paramètres d'un réseau de neurone. Pour réussir un bon apprentissage; il faut réunir de bons ingrédients. En effet, un bon algorithme de minimisation trouve rapidement un minimum; mais celui-ci n'est pas forcément satisfaisant. Le nombre d'exemples d'apprentissage et leur distribution sont des données fondamentales. Ainsi, un grand nombre d'exemples bien distribués garantit une forme plus régulière de la fonction de cout, que ses minimas locaux globaux correspondent bien aux vrai valeurs des paramètres. Le choix d'un algorithme efficace et d'une architecture adaptée garantissent la qualité des résultats. Finalement, nous pouvons affirmer que les RNF se distinguent des autres formes de calcul par deux caractéristiques principales : Ils sont adaptatifs, ils ne sont pas programmés, mais ils sont entraînés à partir de données. De plus, les RNF sont réputés s'améliorer avec l'expérience ; plus ils ont appris de données, plus leur réponse est exacte. A la base de ce qui a été montré dans ces deux premiers chapitres, nous entamons la partie de développement d'un algorithme d'optimisation. Cet algorithme sera basé sur cet outil mathématique qui fera l'objet d'un quatrième chapitre.

3 Chapitre 3 : La thermoélectricité

3.1 Introduction

Notre travail consiste à appliquer les réseaux de neurones à la thermoélectricité afin de développer une technique de caractérisation permettant la mesure simultanée de plusieurs propriétés en même temps donnant ainsi un accès plus fin aux propriétés physiques et donc à la compréhension des systèmes. Plusieurs travaux ont été développés pour la modélisation des phénomènes physiques et chimiques par le modèle neuronal, comme ceux étudiés par notre équipe de recherche [49, 50, 51]. Pour un meilleurs suivi, nous avons exposé les notions de base de l'effet thermoélectrique.

L'effet thermoélectrique est un phénomène physique caractéristique de certains matériaux contribuant à la conversion de l'énergie. Un matériau thermoélectrique permet de transformer directement de la chaleur en électricité (génération d'électricité) ou de déplacer des calories par l'application d'un courant électrique (application de réfrigération). La production d'électricité à partir d'un flux thermique permet d'envisager des dispositifs de production d'électricité "propre" [52]. Il est en effet possible de générer un courant à partir de sources de chaleur perdue comme les pots d'échappement, les circuits de refroidissement des centrales nucléaires,... etc. En dépit de leurs nombreux avantages, les dispositifs thermoélectriques sont, à l'heure actuelle, encore limités à des applications bien spécialisées où leurs mérites permettent de compenser leurs coûts élevés et leur relative faible performance.

Les performances d'un système thermoélectrique sont décrites par le nombre sans dimension zT initialement formalisé par Abram F. Ioffé [53]. Ce paramètre zT , aussi appelé facteur de mérite, est utilisé pour caractériser et comparer les performances des matériaux, il traduit l'impact des paramètres physiques qui sont la conductivité thermique, la conductivité électrique, le coefficient de Seebeck. L'étude complète d'un matériau ou système requiert donc différentes mesures électriques et thermiques avec des conditions aux limites adaptées à chaque type de

mesure, ce qui impose de faire appel à des systèmes expérimentaux différents. La spectroscopie d'impédance électrique (IS) est une technique alternative qui aide à définir les propriétés thermoélectriques des éléments. Le gain majeur de cette technique réside dans son efficacité et sa capacité à caractériser et définir le comportement du procédé électronique à l'interface des matériaux. Le principal avantage de cette méthode par rapport à d'autres outils de caractérisations réside dans la capacité de séparer les processus liés à l'électronique se produisant dans l'appareil. Ceci est réalisé en mesurant la réponse d'impédance du système quand un signal de faible amplitude (tension ou courant) est appliqué. La réponse est en fonction de la fréquence et donc des processus rapides apparaîtront à des fréquences élevées, tandis que les processus lents apparaîtront seulement aux basses fréquences.

Dans cette sous-partie est présentée une introduction à la thermoélectricité permettant de développer les notions de bases de ce domaine, une brève introduction dans l'ordre chronologique des découvertes va nous permettre de comprendre les origines de la thermoélectricité en partant de la découverte fondamentale. Dans un premier temps, les effets thermoélectriques seront illustrés simplement dans le cadre de leur découverte initiale. Dans un second temps, nous nous attarderons sur les démonstrations de ces effets grâce à des modèles simples. Une fois ces bases définies, nous pourrons alors déterminer le rendement de conversion thermoélectrique et dériver la célèbre figure de mérite zT . Nous verrons par la suite l'importance de la caractérisation des matériaux thermoélectriques par la spectroscopie d'impédance.

3.2 Présentation

Un matériau thermoélectrique est capable de transformer directement la chaleur en électricité lorsqu'il est soumis à une différence de température entre ses extrémités (cas générateur), ou de déplacer la chaleur d'un point à l'autre s'il est

traversé par un courant électrique (cas réfrigérateur), tel qu'illustré à la figure (3.1). Cette caractéristique est utilisée dans les modules générateurs thermoélectriques (MGT), pour les réaliser, on connecte deux types de matériaux thermoélectriques, normalement semi-conducteurs soumis à une différence de température entre leurs extrémités, on peut profiter ainsi de la chaleur non utilisée produite dans les procédés industriels, dû à cet avantage, les dispositifs thermoélectriques peuvent être trouvés dans une grande gamme d'applications [54].

En général, un module générateur est un dispositif thermoélectrique constitué

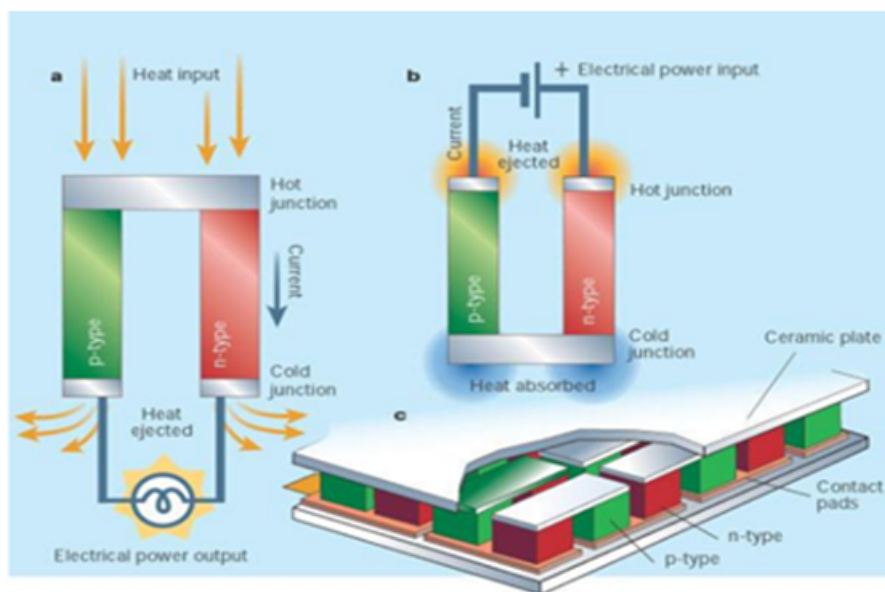


Figure 3.1 – (a) Thermocouple pour générer la puissance électrique (b) Thermocouple pour réfrigérer (c) Dispositif thermoélectrique contenant plusieurs thermocouples [55]

de pattes de matériaux semi-conducteurs (type P et type N) qui sont connectées électriquement en série et thermiquement en parallèle et qui sont exposées à une différence de température pour produire de l'électricité. La performance de ce type de module dépend de la performance des matériaux semi-conducteurs utilisés dans les pattes et de la différence de température imposée, ainsi que d'autres facteurs moins importants liés à la géométrie.

Si nous connectons une source de courant extérieure (une pile par exemple) aux bornes du dispositif (dans le sens indiqué (3.1.a)), une différence de température apparait aux extrémités du thermocouple, de la chaleur est donc absorbée ou dégagée au niveau des jonctions lorsqu'un courant électrique traverse le circuit. Il est ainsi possible d'abaisser ou d'élever la température d'un objet si celui-ci est en contact avec les zones froides ou chaudes. On va donc assimiler notre thermocouple à un réfrigérateur ou une pompe à chaleur dans lesquels les porteurs de charge jouent le rôle du fluide frigorigène (3.1.b). Si plusieurs thermocouples sont disposés électriquement en série, on obtient alors un module thermoélectrique [56] comme celui présenté à la figure (3.2).

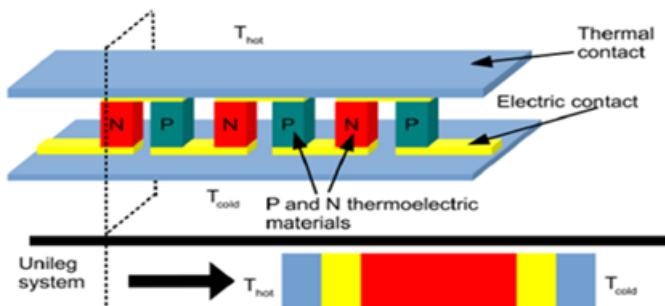


Figure 3.2 – Modèle unidimensionnel du module Peltier

Les modules commerciaux actuels sont fabriqués à partir de matériaux semi-conducteurs tels que $(Bi, Sb)_2(Te, Se)_3$ ou PbTe dont les propriétés thermoélectriques sont optimisées dans une gamme de température donnée [57]. Néanmoins, les performances de ces matériaux en terme de conversion d'énergie restent faibles si bien que le rendement d'un module thermoélectrique plafonne à l'heure actuelle à un pourcentage entre 5 - 7 pourcent. L'optimisation des matériaux reste un défi à relever et justifie les efforts déployés pour découvrir de nouveaux matériaux thermoélectriques encore plus performants [58].

3.3 Effets thermoélectriques

3.3.1 L'effet Seebeck

Il consiste en la génération d'une tension ΔV lorsque les jonctions de deux matériaux conducteurs différents A et B (voir figure 3.3) sont maintenues aux points 1 et 2 à des températures différentes T_1 et T_2 avec $T_1 > T_2$, il se forme alors une différence de potentiel V entre les points 3 et 4, tel que

$$\Delta V = (S_A - S_B) \times (T_1 - T_2) = S_{AB} \times \Delta T \quad (3.1)$$

Où S_{AB} représente le coefficient Seebeck différentiel [59] entre les éléments A et B

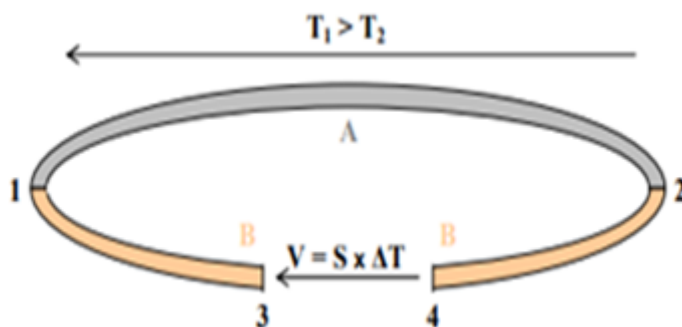


Figure 3.3 – Illustration de l'effet Seebeck [60]

3.3.2 L'effet Peltier

Réciproquement, dans le cas de l'effet Peltier[61], un courant électrique I est imposé au circuit, ce qui entraîne une libération de chaleur Q à une jonction et une absorption de chaleur à l'autre jonction (voir figure 3.4). Le coefficient Peltier Π_{AB} est alors défini par

$$dQ_s = (\Pi_A - \Pi_B) j \vec{U}_{AB}, \quad (3.2)$$

Où dQ_S représente la quantité de chaleur émise (comptée en positif) ou absorbée (comptée en négatif) par unité de surface à l'une des deux jonctions (en $W.m^{-2}$), Π_A et Π_B sont respectivement les coefficients de Peltier des matériaux A et B, le vecteur \vec{j} est le flux de charge (ou la densité surfacique de courant) (en $W.m^{-2}$), et \vec{U}_{AB} le vecteur unitaire adimensionnel dirigé du matériau A vers le matériau B au niveau de la jonction considérée.

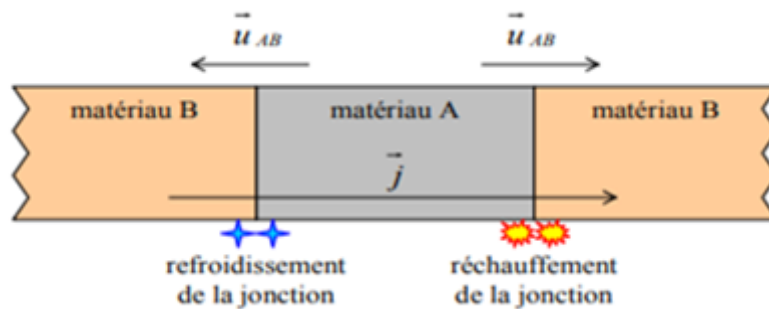


Figure 3.4 – Illustration de l'effet Peltier [60]

Ainsi, le passage d'une densité de courant dans le circuit provoque le refroidissement de la jonction où le courant circule du matériau dont le coefficient Peltier est le plus faible vers le matériau dont le coefficient Peltier est le plus fort (dans le cas de la figure (3.4), $\Pi_A > \Pi_B$) et provoque l'échauffement de l'autre jonction. Ce phénomène est à la base de la réfrigération par effet Peltier.

3.3.3 L'effet Thomson

L'effet Thomson est dû à la variation du coefficient Seebeck en fonction de la température. Si un conducteur est soumis à un gradient de température et est traversé par une densité de courant \vec{j} alors une création ou une absorption de chaleur comme dans l'effet Peltier. C'est une manifestation continue de l'effet Peltier et la chaleur produite est différente de celle produite par l'effet Joule. La quantité de chaleur par unité de temps et par unité de volume dQ_v émise ou

absorbée est proportionnelle au produit scalaire de la densité de courant et du gradient de température [62]).

On peut transcrire ce phénomène sous la forme d'une équation locale

$$dQ_v = -\beta \frac{\partial T}{\partial r} \vec{j}, \quad (3.3)$$

Où dQ_v représente la quantité de chaleur émise (comptée en positif) ou absorbée (comptée en négatif) par unité de volume du matériau (en $W.m^{-3}$), β le coefficient de Thomson (en $V.K^{-1}$), $T(r)$ est la température au point r du matériau (en K), et le vecteur \vec{j} est le flux de charge (en $A.m^{-2}$).

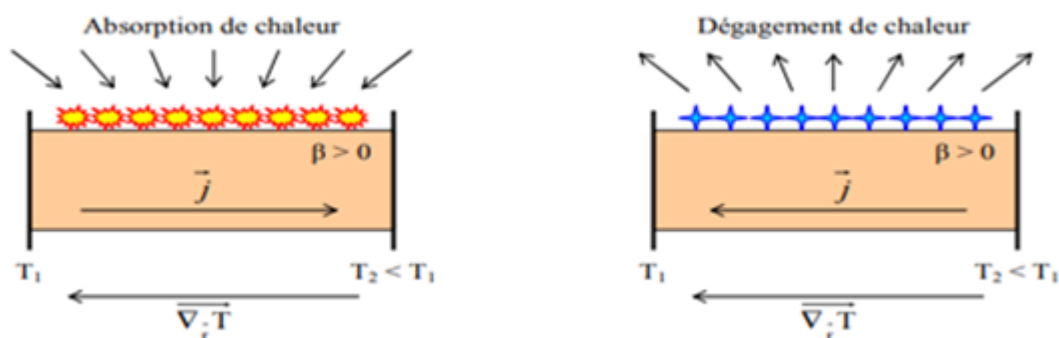


Figure 3.5 – Illustration de l'effet Thomson [60]

3.3.4 Les relations de Kelvin

Les trois coefficients thermoélectriques définis précédemment sont reliés par les deux relations de Kelvin suivantes :

$$S_{AB} = \frac{\Pi_{AB}}{T}, \quad (3.4)$$

$$\frac{\partial S_{AB}}{\partial T} = \frac{\beta_A - \beta_B}{T}, \quad (3.5)$$

L'équation (3.4) démontre la réversibilité de deux effets, Seebeck et Peltier, et permet de déterminer le coefficient Peltier via le coefficient Seebeck. En ce qui concerne l'équation (3.5), elle relie les effets Seebeck et Thomson et permet de définir la valeur absolue du coefficient Seebeck d'un seul matériau selon

$$S_{AB} = \int_0^T \frac{\beta_A}{T} dT - \int_0^T \frac{\beta_B}{T} dT = S_A - S_B, \quad (3.6)$$

D'où le coefficient de Seebeck absolu d'un matériau

$$S = \int_0^T \frac{\beta}{T} dT. \quad (3.7)$$

De la même manière, le coefficient Peltier absolu d'un matériau peut être déduit en utilisant la relation (3.6). Les relations de Kelvin montrent que les trois effets thermoélectriques existent simultanément dans tout matériau thermoélectrique lorsque celui-ci est parcouru par un courant électrique et soumis à une différence de température.

3.4 Conversion d'énergie

Les dispositifs thermoélectriques peuvent être comparés entre eux au travers de différents critères, qui permettent de définir leurs performances ou leurs rendements. Plusieurs critères sont présentés dans cette section. Les dispositifs thermoélectriques, qu'ils soient destinés à la réfrigération ou à la production d'électricité, possèdent le même type d'architecture. Les modules qui composent ces dispositifs sont un assemblage de couples connectés en parallèle du point de vue thermique, et en série du point de vue électrique. Un couple est formé d'un matériau semi-conducteur de type n ($S < 0$) et d'un matériau semi-conducteur de type p ($S > 0$). Ces matériaux composent les deux "jambes" du couple et sont reliés entre eux par un troisième matériau, conducteur, dont le pouvoir thermoélectrique est supposé nul. La figure 3.8 ci-dessous présente un schéma de principe de la génération

d'électricité (gauche) et la réfrigération (droite) dans le cas d'un convertisseur élémentaire, c'est-à-dire un module constitué uniquement d'un couple.

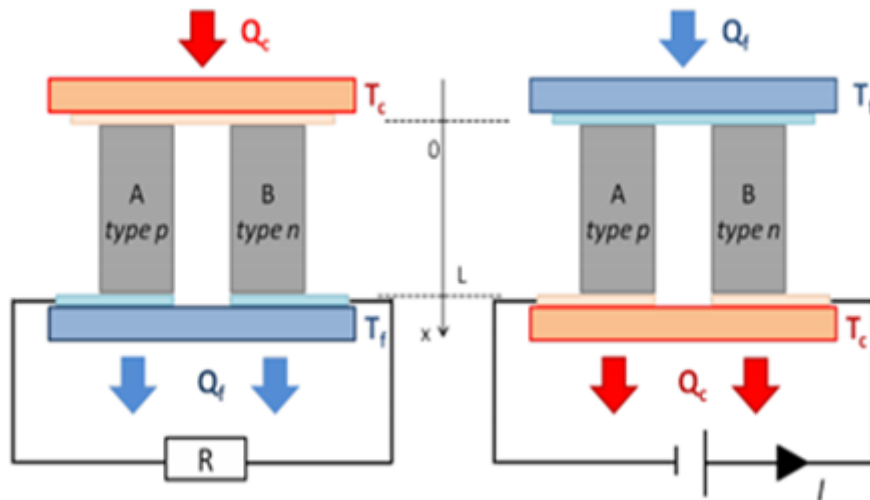


Figure 3.6 – Schéma d'un module unitaire configuré pour la génération d'électricité (gauche) et réfrigération (droite) [63].

Dans le cas de la réfrigération, un courant électrique est appliqué à partir d'une source de courant extérieure, déclenchant un flux de porteurs de charges de la source froide vers la source chaude. Ce déplacement des porteurs de charges contribue à un transfert d'entropie entre les deux sources, et donc à la création d'un flux thermique, qui va s'opposer à celui des porteurs de charges. Ce flux va ainsi entraîner l'abaissement de la température au niveau de la source froide. Il est alors possible de refroidir un objet, si celui-ci est en contact avec les zones froides. Dans le mode génération d'électricité, les branches p et n du couple sont connectées à une charge extérieure et leurs extrémités sont mises en contact avec une source froide et une source chaude. La différence de température aux bornes du couple entraîne le déplacement des porteurs de charges et produit un courant électrique dans le circuit. Une puissance électrique est ainsi dissipée dans la résistance de charge.

3.5 Rendement

Le rendement de conversion thermique-électrique d'un thermocouple est donné par la relation (3.8) qui suppose le système isolé, c'est-à-dire la perte de chaleur négligeable

$$\eta = \frac{P}{Q} = \frac{\text{puissance fournie à la charge}}{\text{chaleur absorbée à l'extrémité chaude}}, \quad (3.8)$$

A partir de cette équation il a été montré [64] que le rendement maximum obtenu est donné par

$$\eta_{\max} = \eta_c \frac{\sqrt{1 + z_{AB} \frac{T_{ch} + T_{fr}}{2}} - 1}{\sqrt{1 + Z_{AB} \frac{T_{ch} + T_{fr}}{2} + \frac{T_{fr}}{T_{ch}}}}, \quad (3.9)$$

η_c est le rendement de carnot caractéristique aux machines thermiques

$$\eta_c = \frac{T_{ch} - T_{fr}}{T_{ch}}, \quad (3.10)$$

Le coefficient Z_{AB} est défini par

$$Z_{AB} = \frac{(S_A - S_B)^2}{\left(\sqrt{\frac{K_A}{\sigma_A}} + \sqrt{\frac{K_B}{\sigma_B}}\right)^2}. \quad (3.11)$$

Cette équation représente aussi la dépendance de l'efficacité en fonction des propriétés intrinsèque des matériaux. Le zT , appelé figure de mérite dépend de $(S, \sigma, K)A, B$ qui représentent respectivement le coefficient Seebeck, la conductivité électrique et la conductivité thermique des matériaux A et B. Le facteur de mérite permet d'évaluer la qualité d'un matériau thermoélectrique en combinant ses trois propriétés.

Les dispositifs thermoélectriques sont actuellement utilisés en réfrigération sous la forme de cellules à effet Peltier semblables à celle présentée à la figure (3.2). La performance d'un matériau thermoélectrique est généralement qualifiée par une figure de mérite, le coefficient zT , qui met en relation les trois propriétés

thermoélectriques fondamentales. Pour un composé comme le Bi_2Te_3 , sa valeur peut varier entre 0.2 et 1.5 selon la température et le dopage [65], alors que la demande actuelle requiert un coefficient qui s'approche de 2. Certains chercheurs dans le domaine de la thermoélectricité croient qu'un coefficient zT rejoignant une valeur rapprochée de 3 sera plausiblement atteignable dans le futur [66].

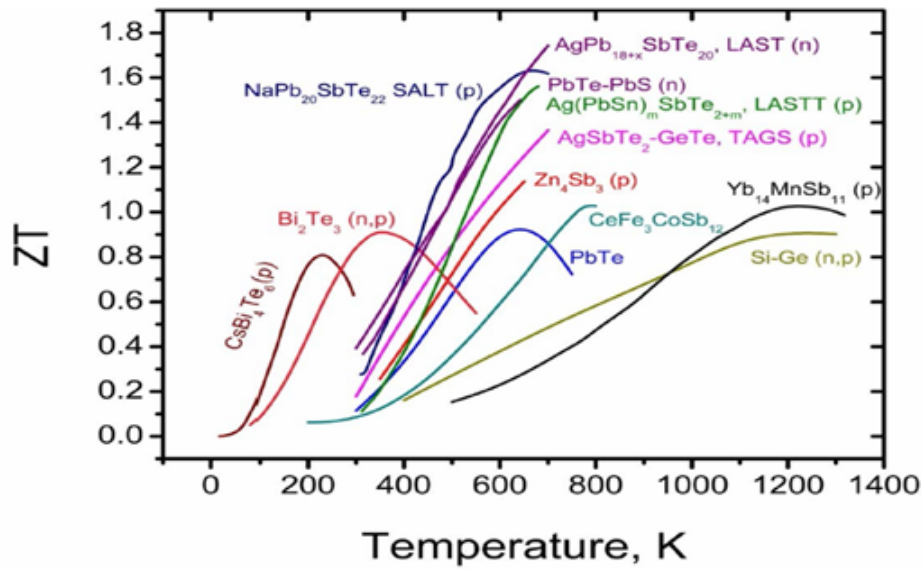


Figure 3.7 – Le facteur de mérite de différents matériaux thermoélectriques en fonction de la température [67].

3.6 Technique de mesure des propriétés thermoélectriques

Il est possible d'envisager une approche de caractérisation dans le cadre de la réponse électrique linéaire dénommée spectroscopie d'impédance. Elle est couramment employée dans divers domaines de recherche, en particulier pour étudier les systèmes électrochimiques, c'est une méthode utile pour la caractérisation des modules thermoélectriques TE. Ça peut déterminer avec une grande précision la figure de mérite zT du module ainsi que les propriétés physiques des thermoéléments par une étude dynamique du système au moyen de sa réponse en fonction de la fréquence d'excitation. C'est une méthode de caractérisation entièrement électrique menée de manière isotherme ou adiabatique, qui ne requiert donc pas de contrôler le gradient de température dans le système, ce qui constitue un atout majeur dans le cadre de l'étude des systèmes thermoélectriques [68]. Pour ces raisons nous avons choisi de reprendre les résultats obtenus par la caractérisation par spectroscopie d'impédance [69, 70] afin de faire une comparaison avec nos résultats obtenus par modélisation neuronale, l'interprétation des résultats d'impédance nécessite généralement l'utilisation d'un modèle théorique (circuit équivalent) qui fournit les paramètres de dispositif souhaités après un ajustement aux résultats expérimentaux en développant un circuit équivalent pour les éléments thermoélectriques suspendus dans des conditions adiabatiques.

3.6.1 La réponse thermique d'un dispositif thermoélectrique

Dans un matériau thermoélectrique (TE), différents processus se produisent lorsqu'un courant constant est appliqué à travers lui. Premièrement, une différence de potentiel initiale est établie à travers l'élément par un champ électrique externe qui produit le courant. Ensuite, en raison de l'effet Peltier, la chaleur est libérée

à l'une des jonctions et absorbée à l'autre, créant progressivement une différence de température à travers l'élément. Cette différence de température sépare les supports dans le matériau, ce qui induit une différence de potentiel Seebeck due à l'effet Seebeck. Cette tension Seebeck ralentit le flux des porteurs et diminue le courant, augmentant ainsi la résistance du système. Tous ces processus ont des échelles de temps distinctes et sont régis par différents paramètres. En utilisant la spectroscopie d'impédance, qui utilise des ondes d'excitation de courant alternatif, nous pouvons extraire des informations très utiles sur la dynamique de ces processus thermoélectriques fondamentaux, qui fournissent une connaissance détaillée unique du fonctionnement du système.

3.6.2 Calcul de la fonction de l'impédance dans le domaine des fréquences

Soit un thermoélément de section transversale A et de longueur L lié à des contacts métalliques très minces L_M . L'élément est suspendu dans des conditions adiabatiques (aucun échange de chaleur avec l'environnement) à une température initiale homogène T_i . Une onde de courant sinusoïdale de faible amplitude de l'ordre de 0 A est utilisée comme signal d'excitation, à cause de la faible amplitude du courant et de la conductivité électrique généralement élevée des matériaux thermoélectriques, l'effet Joule est négligé. De plus, toutes les propriétés thermoélectriques sont supposées indépendantes de la température. La figure (3.8) montre les différents processus thermiques qui apparaissent dans un thermoélément de type n sous ces hypothèses. Il est à noter qu'un plan à température constante est présent à la demi-longueur $L_H = L/2$, où la température reste à sa valeur initiale à tout moment [69].

Les flèches indiquent dans la figure (3.8) la direction des flux de chaleur conducteurs apparaissant aux jonctions en conséquence de l'effet Peltier quand un

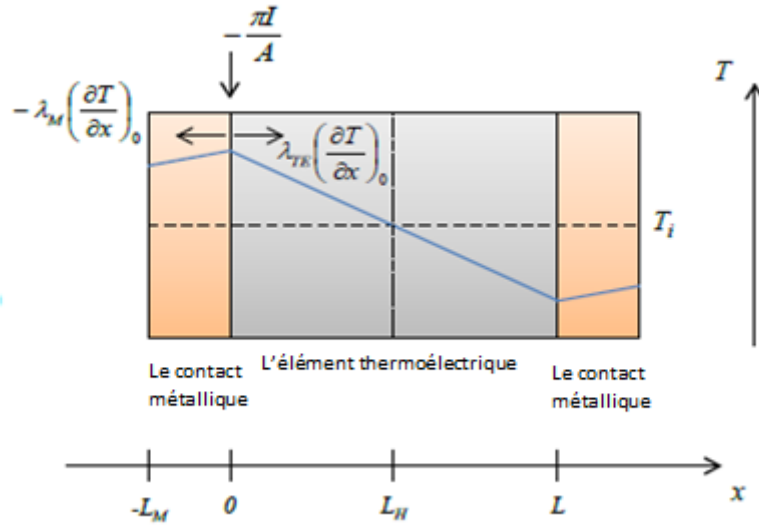


Figure 3.8 – Modèle physique d'un élément thermoélectrique lié à deux contacts métalliques[69].

matériau thermoélectrique de type n et un courant positif sont considérés. L'impédance est donnée par le rapport de différence de potentiel V à travers l'élément et du courant I circulant à travers le circuit externe, la chute de tension du système est causée par deux contributions, la première en raison de la résistance ohmique R du système (qui inclut les résistances électriques intrinsèques du matériau R_i et les contacts métalliques R_m et l'autre est produite par l'effet Seebeck. Prendre ceci en compte, la fonction d'impédance devient:

$$Z = R + \frac{S[T(L) - T(0)]}{1}, \quad (3.12)$$

Où $T(0)$ et $T(L)$ sont les températures à $x=0$ et $x=L$, respectivement. L'analyse néglige l'effet Seebeck dans les contacts puisque les métaux présentent généralement un S beaucoup plus petit que celui des matériaux thermoélectriques et il est limité aux mesures d'impédance effectuées sous une petite perturbation AC autour d'un état initiale d'équilibre thermique stationnaire. Il ressort de l'équation

(3.12) que pour identifier la fonction de l'impédance, il est nécessaire de connaître l'évolution avec la fréquence des températures sur les bords de l'élément ($T(L) - T(0)$). Cela implique de résoudre l'équation de la chaleur du système.

3.6.3 Influence du contact métallique

a-Lorsque l'on néglige l'influence des contacts métalliques

Un processus de diffusion thermique est un phénomène de transfert thermique irréversible qui se produit sans modifier le mouvement macroscopique du système. Ce transfert se produit dans un système initialement hors d'équilibre, des régions chaudes aux régions froides; dans une première approche, la présence des contacts était négligée (hypothèse de contacts extrêmement minces), en réalité, des contacts sont toujours nécessaires, sinon aucun effet Peltier ne sera produit.

L'équation de la chaleur du système (lorsque l'on néglige les contacts métalliques) à $0 < x < L$ est donnée par :

$$\frac{\partial^2 T}{\partial x^2} = \frac{1}{\alpha_{TE}} \frac{\partial T}{\partial t}, \quad (3.13)$$

où α_{TE} est la diffusivité thermique de l'élément thermoélectrique, qui est liée à la conductivité thermique (λ), la densité de masse (ρ), et la chaleur spécifique (C_p) de l'élément thermoélectrique

$$\alpha_{TE} = \frac{\lambda_{TE}}{\rho \cdot C_p}, \quad (3.14)$$

si $\Delta T = T - T_i$ est considérée comme la différence de température par rapport à T_i alors.

$$\partial \Delta T = \partial T, \quad (3.15)$$

définissons la fonction θ la transformée de Laplace [71] de la différence de tem-

pérature par rapport à la température initiale

$$\theta(x, j\omega) = L (T(x, t) - T_i) \quad (3.16)$$

t est le temps et x est la position.

$$\theta = L(\Delta T) = \int_0^\infty \Delta T \exp(-j\omega t) dt \quad (3.17)$$

L'équation (3.13) est ainsi convertie du domaine temporel au domaine fréquentiel ($j\omega$)

$$\frac{\partial^2 \theta}{\partial x^2} - \frac{j\omega}{\alpha_{TE}} \theta = 0 \quad \text{à} \quad 0 < x < L \quad (3.18)$$

où $j = \sqrt{-1}$ et ω est la fréquence angulaire $\omega = 2\pi f$

f : la fréquence d'oscillation du signal AC.

La solution de cette équation prend la forme

$$\theta(x, j\omega) = C_1 \sinh \left[\frac{x}{L_H} \left(\frac{j\omega}{\omega_{TE}} \right)^{0.5} \right] + C_2 \cosh \left[\frac{x}{L_H} \left(\frac{j\omega}{\omega_{TE}} \right)^{0.5} \right] \quad (3.19)$$

C_1 et C_2 sont des constantes et ω_{TE} est la fréquence angulaire caractéristique qui est liée à la diffusivité thermique et la longueur par

$$\omega_{TE} = \frac{\alpha_{TE}}{L_H^2} \quad (3.20)$$

sa dérivée par rapport à la position est

$$\frac{\partial \theta}{\partial x}(x, j\omega) = \frac{1}{L_H} \left(\frac{j\omega}{\omega_{TE}} \right)^{0.5} \left\{ C_1 \cosh \left[\frac{x}{L_H} \left(\frac{j\omega}{\omega_{TE}} \right)^{0.5} \right] + C_2 \sinh \left[\frac{x}{L_H} \left(\frac{j\omega}{\omega_{TE}} \right)^{0.5} \right] \right\} \quad (3.21)$$

Les conditions aux limites données par la température constante dans le plan ($x = L_H = L/2$) et les limites adiabatiques peuvent également être définis dans le

domaine fréquentiel comme

$$T\left(\frac{L}{2}, t\right) = T_i \rightarrow \theta\left(\frac{L}{2}, \omega\right) \quad (3.22)$$

et

$$\frac{-\pi_0 I_0}{A} + \lambda_{TE} \left(\frac{\partial T}{\partial x}\right)_0 = 0 \rightarrow \frac{-\pi_0 i_0}{A} + \lambda_{TE} \left(\frac{\partial \theta}{\partial x}\right)_0 = 0 \quad (3.23)$$

où π_0 est le coefficient de Peltier

$$\pi_0 = ST_i \quad (3.24)$$

et i_0 est la transformée de Laplace du courant I_0 à $x = 0$.

$$L[I(0, t)] = i_0 \quad (3.25)$$

En prenant $T(0) \approx T_i$ due au faible signal d'excitation AC appliqué. (des grands changements ne sont pas attendus dans la température de la jonction)

En appliquant ces conditions aux limites aux équations (3.19) et (3.21), les constantes d'intégration peuvent être obtenues

$$C_1 = \frac{i_0 \pi_0 L_H}{\lambda_{TE} A} \left(\frac{j\omega}{\omega_{TE}}\right)^{-0.5} \quad (3.26)$$

$$C_2 = -\frac{i_0 \pi_0 L_H}{\lambda_{TE} A} \left(\frac{j\omega}{\omega_{TE}}\right)^{-0.5} \tanh \left\{ \left(\frac{j\omega}{\omega_{TE}}\right)^{0.5} \right\} \quad (3.27)$$

On obtient alors

$$\theta(x, j\omega) = \frac{ST_i i_0 L_H}{\lambda_{TE} A} \left(\frac{j\omega}{\omega_{TE}}\right)^{-0.5} \left\{ \sinh \left[\frac{x}{L_H} \left(\frac{j\omega}{\omega_{TE}}\right)^{0.5} \right] - \cosh \left[\frac{x}{L_H} \left(\frac{j\omega}{\omega_{TE}}\right)^{0.5} \right] \tanh \left[\left(\frac{j\omega}{\omega_{TE}}\right)^{0.5} \right] \right\} \quad (3.28)$$

$$\theta_0 = \theta(0, j\omega) = - \left[\left(\frac{sT_i i_0 L_H}{\lambda_{TE} A} \right) \right] \left(\frac{j\omega}{\omega_{TE}} \right)^{-0.5} \tanh \left[\left(\frac{j\omega}{\omega_{TE}} \right)^{0.5} \right] \quad (3.29)$$

Puisque la différence de potentiel Seebeck

$$V_s = V_s(0) - V_s(L) = -S[T(0) - T(L)] = -2 S [T(0) - T_i] \quad (3.30)$$

à cause de la présence du plan à température constante.

Et comme

$$T(L) - T(0) = -2\theta(0) \quad (3.31)$$

donnée dans le domaine des fréquences

Et la différence de potentiel totale $V = V(0) - V(L)$ inclut la chute de tension ohmique IR

$$V = IR - 2 S [T(0) - T_i] \quad (3.32)$$

La fonction de l'impédance dans le domaine des fréquences peut être obtenue par

$$Z(j\omega) = R - \frac{2 S \theta(0, j\omega)}{i_0} = R + \frac{S^2 T_i L}{\lambda_{TE} A} \left(\frac{j\omega}{\omega_{TE}} \right)^{-0.5} \tanh \left\{ \left(\frac{j\omega}{\omega_{TE}} \right)^{0.5} \right\} \quad (3.33)$$

On peut observer que le terme de la résistance thermoélectrique $R_{TE} = \frac{s^2 T_i L}{\lambda_{TE} A}$ apparait dans l'expression de Z .

Le deuxième terme de l'équation (3.33) prend la forme d'un Warburg à température constante ω_{CT}

$$Z_{\omega_{CT}}(j\omega) = R_{TE} \left(\frac{j\omega}{\omega_{TE}} \right)^{-0.5} \tanh \left\{ \left(\frac{j\omega}{\omega_{TE}} \right)^{0.5} \right\} \quad (3.34)$$

Cette dénomination est une conséquence de la similitude avec la longueur finie de Warburg et la frontière transmissive (court Warburg) élément fréquemment utilisé

en électrochimie, qui est dans le modèle de Poisson-Nernst-Planck.

b- Influence des contacts métalliques

La présence du contact métallique dans un dispositif thermoélectrique permet à la chaleur Peltier générée à chaque jonction de pouvoir diffuser à travers son épaisseur (voir figure 3.8) en dehors de la diffusion vers la demi-longueur des thermoéléments. Cela influencera considérablement la réponse d'impédance. Afin d'évaluer l'influence des contacts métalliques. Il est nécessaire de résoudre l'équation de la chaleur en considérant leur influence.

$$\frac{\partial^2 T}{\partial x^2} = \frac{1}{\alpha_M} \frac{\partial T}{\partial t} \quad -L_M < x < 0 \quad (3.35)$$

En suivant la même procédure que ci-dessus, on peut obtenir dans le domaine fréquentiel

$$\frac{\partial^2 \theta}{\partial x^2} - \frac{j\omega}{\alpha_M} \theta = 0 \quad \text{à} \quad -L_M < x < 0 \quad (3.36)$$

α_M est la diffusivité du métal qui est liée à la conductivité thermique λ_M , la densité de masse d_M et la chaleur spécifique du contact C_{PM} , $\alpha_M = \frac{\lambda_M}{d_M C_{PM}}$. Les conditions aux limites sont

$$\left(\frac{\partial \theta}{\partial x} \right)_{-L_M} = 0 \quad \text{à} \quad x = -L_M \quad (3.37)$$

$$-\frac{i_0 \pi_0}{A} - \lambda_M \left(\frac{\partial \theta}{\partial x} \right)_{0,M} + \lambda_{TE} \left(\frac{\partial \theta}{\partial x} \right)_{0,TE} = 0 \quad \text{à} \quad x = 0 \quad (3.38)$$

$$\theta(0, j\omega)_M = \theta(0, j\omega)_{TE} \quad \text{à} \quad x = 0 \quad (3.39)$$

$$\theta(L_H, j\omega) = 0 \quad \text{à} \quad x = L_H \quad (3.40)$$

L'équation (3.38) assure la continuité de la température à $x = 0$, après la résolution du système d'équations on obtient, pour $-L_M \leq x \leq 0$

$$\theta(x, j\omega) = -\frac{sT_i i_0}{A} \frac{\left\{ \tanh \left[\left(\frac{j\omega}{\omega_M} \right)^{0.5} \right] \sinh \left[\frac{x}{L_M} \left(\frac{j\omega}{\omega_M} \right)^{0.5} \right] + \cosh \left[\left(\frac{x}{L_H} \left(\frac{j\omega}{\omega_M} \right)^{0.5} \right) \right] \right\}}{\left\{ \frac{L_M}{\lambda_M} \left(\frac{j\omega}{\omega_M} \right)^{-0.5} \coth \left[\left(\frac{j\omega}{\omega_M} \right)^{0.5} \right] \right\}^{-1} + \left\{ \frac{L_H}{\lambda_{TE}} \left(\frac{j\omega}{\omega_{TE}} \right)^{-0.5} \tanh \left[\left(\frac{j\omega}{\omega_{TE}} \right)^{0.5} \right] \right\}^{-1}} \quad (3.41)$$

Pour $0 \leq x \leq L$

$$\theta(x, j\omega) = \frac{sT_i i_0}{A} \frac{\left\{ \coth \left(\frac{j\omega}{\omega_{TE}} \right)^{0.5} \right\} \sinh \left[\frac{x}{L_H} \left(\frac{j\omega}{\omega_{TE}} \right)^{0.5} \right] - \cosh \left[\frac{x}{L_H} \left(\frac{j\omega}{\omega_M} \right)^{0.5} \right]}{\left\{ \frac{L_M}{\lambda_M} \left(\frac{j\omega}{\omega_M} \right)^{-0.5} \coth \left[\left(\frac{j\omega}{\omega_M} \right)^{0.5} \right] \right\}^{-1} + \left\{ \frac{L_H}{\lambda_{TE}} \left(\frac{j\omega}{\omega_{TE}} \right)^{-0.5} \tanh \left[\left(\frac{j\omega}{\omega_{TE}} \right)^{0.5} \right] \right\}^{-1}} \quad (3.42)$$

à $x = 0$ les deux équations sont égales et prennent la forme

$$\theta(0, j\omega) = -\frac{sT_i i_0}{A} \left\{ \left\{ \frac{L_M}{\lambda_M} \left(\frac{j\omega}{\omega_M} \right)^{-0.5} \coth \left[\left(\frac{j\omega}{\omega_M} \right)^{0.5} \right] \right\}^{-1} + \left\{ \frac{L_H}{\lambda_{TE}} \left(\frac{j\omega}{\omega_{TE}} \right)^{-0.5} \tanh \left[\left(\frac{j\omega}{\omega_{TE}} \right)^{0.5} \right] \right\}^{-1} \right\} \quad (3.43)$$

Qui peut s'écrire sous la forme

$$\theta(0) = -\frac{\pi_0 i_0}{A} \left(\frac{1}{z_{th,TE}^{-1} + Z_{th,M}^{-1}} \right) \quad (3.44)$$

La différence de potentiel total est

$$V = IR - 2 S [T(0) - T_i] \quad (3.45)$$

On en déduit alors l'expression de l'impédance

$$Z(j\omega) = R + 2 \frac{S^2 T_i}{A} \left\{ \left\{ \frac{L_M}{\lambda_M} \left(\frac{j\omega}{\omega_M} \right)^{-0.5} \coth \left[\left(\frac{j\omega}{\omega_M} \right)^{0.5} \right] \right\}^{-1} + \left\{ \frac{L_H}{\lambda_{TE}} \left(\frac{j\omega}{\omega_{TE}} \right)^{-0.5} \tanh \left[\left(\frac{j\omega}{\omega_{TE}} \right)^{0.5} \right] \right\}^{-1} \right\} \quad (3.46)$$

Cette impédance se consiste en une pure résistance R connecté en série avec deux

éléments Warburg connectés en parallèle ($Z_{W_{CT}}$ et Z_{W_a}).

$$Z = R + \frac{1}{Z_{W_{CT}}^{-1} + Z_{W_a}^{-1}} \quad (3.47)$$

Le circuit équivalent est composé d'une résistance liée en série avec deux Warburg liés parallèlement $Z_{W_{CT}}^{-1}$ et $Z_{W_a}^{-1}$

$$Z_{W_{CT}}(j\omega) = R_{TE} \left(\frac{j\omega}{\omega_{TE}} \right)^{-0.5} \tanh \left\{ \left(\frac{j\omega}{\omega_{TE}} \right)^{0.5} \right\} \quad (3.48)$$

$$Z_{W_a} = R_M \left(\frac{j\omega}{\omega_M} \right)^{-0.5} \coth \left\{ \left(\frac{j\omega}{\omega_M} \right)^{0.5} \right\} \quad (3.49)$$

Z_{W_a} est définie comme d'un Warburg adiabatique W_a , qui donne une expression similaire à un Warburg de longueur finie avec blocage d'éléments frontaliers (Warburg ouvert) du modèle de Poisson-Nernst-Planck.

La spectroscopie d'impédance est analysée en tant qu'outil de caractérisation de matériaux et dispositifs thermoélectriques par le développement de circuits équivalents appropriés fournissant un excellent ajustement à la réponse d'impédance des processus thermoélectriques. En utilisant le circuit équivalent des équations (3.47) et (3.33) plusieurs paramètres clés peuvent être extraits d'une mesure d'impédance relativement simple. Cela fournit potentiellement une technique efficace et rapide pour la caractérisation des propriétés thermoélectriques des systèmes. Le potentiel de cette méthode, contrairement aux expériences réalisées dans le domaine temporel, qui sont généralement difficiles à analyser avec des réponses couramment régies par des processus qui se chevauchent, réside dans leur haute sensibilité et leur capacité à séparer les différents processus physiques se produisant dans les appareils ce qui permet la séparation et la détermination directe des différents processus se produisant dans les appareils et dans les conditions réelles de fonctionnement. Les résultats de cette technique [67, 68] seront exposés dans le chapitre suivant.

3.7 Conclusion

Dans ce chapitre, nous avons présenté les principaux effets thermoélectriques, et les grandeurs physiques principalement intervenant tel que le rendement de conversion et le facteur de mérite zT en effectuant le bilan énergétique d'une machine thermique. L'optimisation intrinsèque des matériaux thermoélectriques peut se réduire uniquement à l'optimisation du zT . De plus, c'est un paramètre sans dimension qui fait intervenir directement les propriétés physiques. Nous avons ensuite abordé le calcul de l'impédance qui s'obtient en résolvant l'équation de la chaleur et qui permet de montrer l'intérêt de la caractérisation par spectroscopie d'impédance. En effet, la spectroscopie d'impédance est une technique prometteuse pour caractériser les systèmes thermoélectriques. En séparant les réponses thermiques et électriques dans le domaine fréquentiel, la mesure par spectroscopie d'impédance fournit tous les paramètres clés pour caractériser les systèmes thermoélectriques.

**4 Chapitre 4 : Les réseaux
de neurones pour prédire
les propriétés thermoélec-
triques des éléments**

4.1 Introduction

Divers domaines de la science ont utilisé les réseaux de neurones artificiels comme outils puissants dans une large gamme d'applications. Ils ont de plus en plus de succès et attirent encore une attention importante pour de nombreuses questions liées à la modélisation des phénomènes physiques et chimiques [72, 76]. En effet, ils restent un sujet de grand intérêt pour les chercheurs qui souhaitent améliorer la performance de ces approches et étendre leur champ d'application [77, 81]. Nous citons, par exemple, leur pertinence dans les solides conducteurs et les nanotechnologies [82, 83], les problèmes de dynamique des fluides [84], la physique des plasmas [85] et la prévision de la corrosion des métaux [86, 87]. Nous nous référons ainsi aux travaux précédents de notre équipe de recherche utilisant les réseaux de neurones artificiels dans la reconnaissance, le traitement de l'information et la prévision des propriétés des matériaux [88, 92].

Le développement de nouvelles techniques et méthodologies expérimentales donne la possibilité d'aller encore plus loin dans l'étude des architectures neuronale en raison de la disponibilité de l'ensemble de données. Ainsi, des modèles plus élaborés sont nécessaires pour étudier et comprendre les mécanismes fondamentaux de nombreux systèmes complexes. Parallèlement à cette évolution de la science et des approches expérimentales, il existe un besoin constant de développement des méthodes mathématiques et nouvelles approches neuronales pour analyser les propriétés de matériaux par exemple.

De nombreuses applications prospectives tentent aujourd'hui d'augmenter les performances des dispositifs thermoélectriques, ces dispositifs sont largement utilisés dans différentes applications telles que la collecte de l'énergie de chaleur perdue dans les nano-dispositifs[93], la production d'électricité pour les dispositifs dans les missions spatiales longues [94]. L'utilisation de matériaux thermoélectriques est

principalement liée à leurs vertus lors des processus de chauffage-refroidissement. Ils sont largement utilisés pour la production d'électricité en raison de leur capacité à convertir le gradient de température en différence de potentiel électrique [95, 96]. Cependant, le développement de meilleurs éléments thermoélectriques reste crucial pour concevoir des dispositifs thermoélectriques efficaces dans différents domaines technologiques.

L'efficacité des matériaux thermoélectriques actuels est encore limitée. La recherche de matériaux plus efficaces est guidée par l'optimisation de trois propriétés la conductivité électrique σ , le coefficient Seebeck S et la conductivité thermique λ (qui est l'addition de la conductivité thermique du réseau et de la conductivité thermique électronique). Celles-ci définissent une valeur sans dimension du mérite $zT = \frac{(\sigma S^2 T)}{\lambda}$. zT est généralement obtenu par la mesure indépendante en fonction de la température de σ , S et λ . Cela nécessite généralement au moins deux instruments différents. S et σ peuvent être mesurés à l'aide d'un seul appareil, tandis que la méthode la plus fréquemment utilisée pour déterminer λ est la technique du flash laser, qui fournit la diffusivité thermique α . Ainsi, la connaissance de la chaleur spécifique C_p et de la masse volumique ρ est nécessaire car $\lambda = \alpha \rho C_p$. La mesure de la chaleur spécifique nécessite une mesure supplémentaire effectuée soit en utilisant le même équipement flash laser ou un autre instrument. Pour la densité de masse, une balance d'Archimède est fréquemment utilisée. Le nombre important d'instruments requis, chacun avec ses propres sources d'erreur, et le grand nombre de mesures à effectuer font de la caractérisation complète de ces matériaux assez longs. De plus, une grande partie de l'équipement requis est assez cher et n'est donc pas toujours facilement accessible à tous les chercheurs. Tous ces inconvénients entraînent des obstacles importants à la recherche de meilleurs matériaux et la réalisation d'une conversion plus efficace de l'énergie thermique en électricité[97].

Le développement de techniques de caractérisation est constamment important pour la mesure simultanée de plusieurs propriétés. En effet, il est particulièrement important de simplifier le protocole de mesure et de réduire les sources d'incertitude. Les travaux menés, dans cette contribution, visent à développer une technique de caractérisation simple et commune capable d'extraire certaines quantités physiques d'éléments thermoélectriques. De plus, il est évident que l'utilisation des ensembles de données disponibles pour la prévision fournit des efforts considérables aux scientifiques qui étudient les propriétés des matériaux dans différents domaines. Tirer parti des approches de réseaux de neurones artificiels et s'appuyer sur l'analyse des corrélations entre l'ensemble de données disponible fournit des informations utiles pour concevoir de nouveaux matériaux dans différentes applications. Pour cette raison, nous avons développé un modèle de réseaux de neurones pour étudier et définir les propriétés thermoélectriques de certains éléments.

4.2 Modélisation des propriétés thermoélectriques des éléments par l'approche des réseaux neuronaux

La partie principale de cette thèse porte sur une étude théorique d'un modèle de réseau de neurones capable de définir et prédire les différents paramètres caractérisant les matériaux thermoélectriques [98]. La topologie du modèle de réseau neuronal est fortement corrélée au problème étudié et aux processus sélectionnés pour traiter complètement les données. Dans une telle approche neuronale, une description adéquate des matériaux étudiés sera prouvée par la possibilité de construire des neurones artificiels capables d'optimiser les données et de calculer les paramètres thermoélectriques à savoir : la conductivité thermoélectrique λ_{TE} , la diffusivité thermoélectrique α_{TE} , le coefficient Seebeck S , la résistance thermoélec-

trique R_{TE} et la fréquence caractéristique de diffusion thermique ω_c estimées par l'analyse de l'impédance équivalente de l'interface étudiée. Pour cela, nous définissons l'impédance électrique de certains éléments comme le Tellure de Bismuth, le Zinc, le Cuivre, l'argent, le cadmium et le thulium, elle est obtenue en résolvant l'équation de la chaleur dans le domaine fréquentiel à savoir

$$Z(j\omega) = R + \frac{S^2 T_i L}{\lambda_{TE} A} \left(\frac{j\omega}{\omega_{TE}} \right)^{-0.5} \tanh \left[\left(\frac{j\omega}{\omega_{TE}} \right)^{0.5} \right] \quad (4.1)$$

notre principale préoccupation est de construire un modèle neuronal capable d'identifier les caractéristiques des éléments thermoélectriques. Les prédictions du réseau neuronal sont ensuite évaluées et comparées aux mesures expérimentales des propriétés thermoélectriques de ces éléments. L'entraînement et l'évaluation du réseau de neurones artificiel impliquent trois phases qui comprennent

- i) la génération de la base de données, qui caractérise l'ensemble d'unités entrée-sortie de la tâche à étudier,
- ii) la construction de la topologie de réseau en fonction des données,
- iii) la vérification de la performance de l'architecture neuronale.

4.2.1 Configuration de la structure du réseau

La définition de l'ensemble de données d'apprentissage est essentielle pour former et définir les paramètres de réseau optimaux. Le modèle neuronal est donc conçu avec une architecture structurale, comprenant la collection d'entrées-sorties, le type de fonction d'activation et les algorithmes d'apprentissage, la qualité et la mise en place de l'ensemble de données du réseau est un point essentiel pour concevoir l'architecture du réseau neuronal. En effet, le contrôle du processus de formation est principalement lié aux unités d'entrée-sortie. À cet effet, les unités d'entrée-sortie sont extraites de l'équation de l'impédance (4.1) et collectés pour former la base de données du modèle du réseau de neurone artificiel. À cette fin,

nous définissons l'entrée par l'expression de l'impédance Z dans les plans réel et imaginaire et les unités de sortie par le coefficient Seebeck et la conductivité thermique. Pour définir les ensembles empiriques nécessaires à la formation du réseau, les valeurs du coefficient Seebeck, de la conductivité thermique et de la fréquence doivent être collectés dans la gamme de leurs valeurs extraites des mesures expérimentales. Sur la base de ces hypothèses, nous formons le modèle du réseau de neurone artificiel pour saisir les paramètres du réseau qui définissent le mieux le modèle optimal en utilisant la technique d'apprentissage supervisé et des algorithmes de second ordre pour mettre à jour les paramètres de pondération de notre réseau neuronal. De plus, la phase de prétraitement est un facteur clé pour fixer le point de départ qui conduit à définir le réseau optimal. Pour cette raison, dans la phase de prétraitement du modèle neuronal, il convient de définir les conditions géométriques de l'échantillon (la longueur L , la surface A), la température (T) ainsi que la densité massique (ρ) de l'élément étudié en tant que paramètres locaux. La qualité de l'ensemble d'apprentissage (exemples constituant les données d'entraînement) et la diversité des valeurs sont les facteurs qui influencent le processus d'apprentissage du réseau. Sur la base des déclarations de données, une représentation schématique du modèle adopté est représentée sur la figure (4.1).

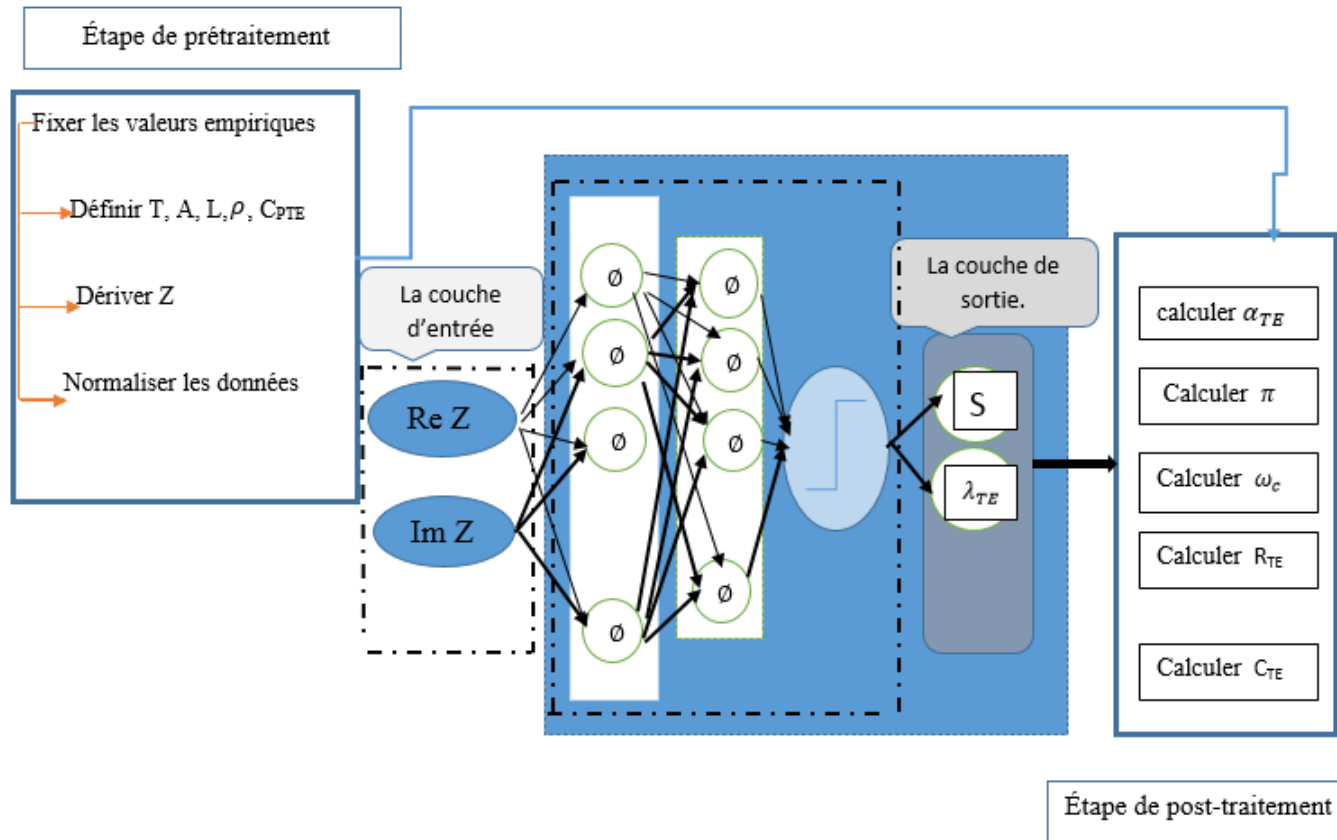


Figure 4.1 – Modèle de réseau neuronal développé pour prédire les propriétés thermoélectriques.

Ce modèle est choisi pour prédire les propriétés thermoélectriques comme le coefficient Seebeck S et la conductivité thermique λ_{TE} . De plus, les autres caractéristiques thermoélectriques sont dérivées de l'étape de post-traitement tel que la capacité thermique C_{TE} le coefficient Peltier Π , la diffusivité thermique α_{TE} et la fréquence caractéristique ω_c et la résistance thermoélectrique R_{TE} .

4.2.2 Mise en place de la phase d'apprentissage du réseau

Le choix d'une topologie neuronale dépend de la complexité du problème, du nombre de couches dans le réseau et du nombre de neurones dans chaque couche et du type des fonctions d'activation. Le bon choix de ces quatre étapes détermine la

capacité du réseau à bien répondre à la tendance des données (tâches à modéliser) [2]. A cette fin, nous construisons une architecture neuronale d'un système thermoélectrique dans des conditions données (température, forme). La représentation schématique de la topologie du réseau est sur la figure (4.2).

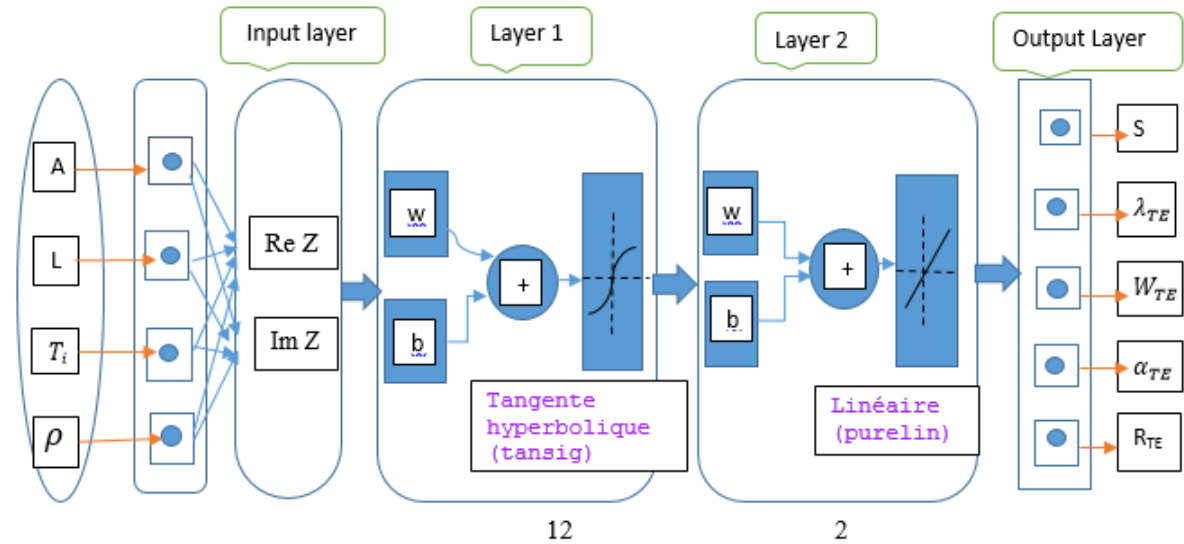


Figure 4.2 – Représentation des différentes couches du réseau de neurone artificiel

La couche de sortie possède les datas des cinq valeurs générées à partir de l'équation (4.1). Ces valeurs sont organisées selon cette topologie. Le nombre de neurones dans la première couche a été traité comme un paramètre variable pour trouver le nombre optimal de neurones adéquat (assurer une précision plus élevée du modèle sans être coincée dans les optima locaux). Les différentes couches de l'architecture donnée ont fait l'objet d'une analyse approfondie. Par exemple, pour illustrer l'impact des unités neuronales, nous avons essayé plusieurs topologies de réseau dans lesquelles le neurone des unités variera de 2 à 12 dans la première couche.

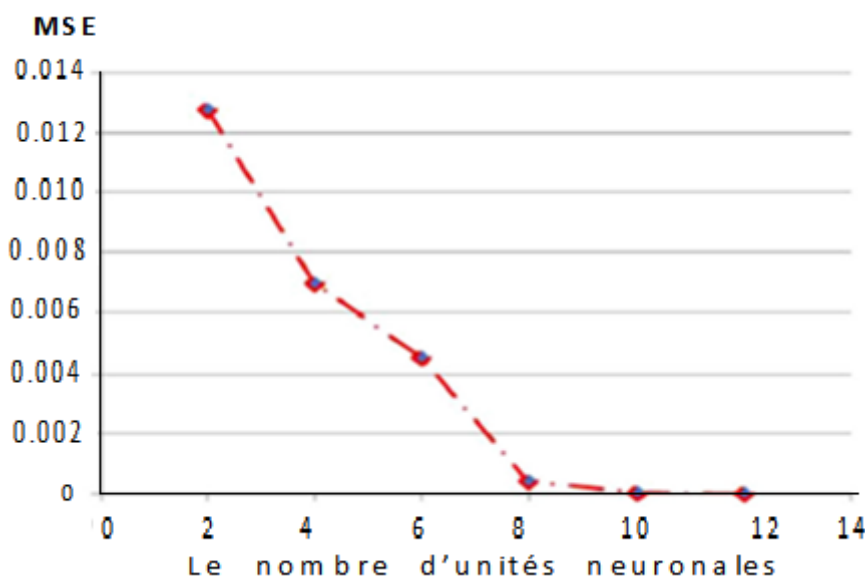


Figure 4.3 – Erreur quadratique moyenne par rapport au nombre d'unités neuronales dans la première couche cachée.

Les résultats illustrent les performances du réseau neuronal en termes de MSE (erreur quadratique moyenne). En fait, les valeurs du MSE montrent qu'une augmentation du nombre de neurones améliorerait les performances de l'approche du réseau. Selon les résultats, le modèle optimal qui capture le mieux le comportement de nos données est donné en utilisant douze neurones dans la première couche cachée.

Le modèle est illustré en traçant le MSE entre la réponse du réseau et les cibles souhaitées à partir des mesures par caractérisation par spectroscopie d'impédance [69, 70]. En fait, le but principal de soumettre le modèle de réseau à différentes architectures est de définir le modèle optimal, qui est capable de fournir les unités d'entrée-sortie dans une approche particulière. Cette étape est très importante, car un mauvais choix de topologie de réseau conduit à un modèle neuronal où sa réponse en sortie serait effectivement sur-paramétrisée. Au final, le modèle optimal a été obtenu par la topologie dans laquelle le nombre de neurones était alimenté par

douze neurones dans la première couche cachée et deux neurones dans la couche de sortie du modèle.

De plus, il existe d'autres facteurs qui peuvent également affecter les performances des réseaux de neurones, par exemple, la fonction d'activation de chaque noeud dans les couches du réseau. Pour notre modèle on a choisi la fonction tangente hyperbolique car elle permettait d'obtenir une meilleure précision. De plus, il existe un autre facteur qui influence la phase d'apprentissage du modèle de réseau neuronal. En fait, la multiplicité des entrées-sorties de la formation joue un rôle clé pour former efficacement l'approche neuronale. Si la diversité de l'ensemble de données est bien choisie, les performances du réseau neuronal seront plus précises.

4.2.3 Validation du modèle par représentation de Nyquist

Nous présentons les résultats de simulation de notre approche adoptée en illustrant la performance du modèle neuronale pour prédire les propriétés thermoélectriques de certains éléments. Les résultats en termes d'impédance présentés dans le plan complexe sont également connus sous le nom de diagramme de Nyquist. Dans cette représentation, la composante imaginaire Im de l'impédance est représentée en fonction de sa partie réelle Re pour une gamme de fréquences d'excitation de 1 mHz à 10 MHz.

Le réseau de neurone artificiel doit définir les paramétrisations de prétraitement qui définissent les paramètres du réseau local liés à la condition de démarrage de l'échantillon. Pour cela, nous avons défini les paramètres locaux du réseau (la géométrie, la chaleur spécifique (C_P), la température initiale) en fonction des expériences données réalisées par Jorge Garcia et al [69],[70] sur le tellure de Bismuth Bi_2Te_3 dans trois cas différents

1. Dans le premier cas le Bi_2Te_3 est caractérisé par une surface $A = 1.4 \times 1.4\text{mm}^2$, et une longueur $L = 1.6\text{mm}$ de longueur figure (4.4).

2. Le second est caractérisé par une surface de $1mm^2$ et une longueur de $1.5mm$ figure (4.5).
3. Le troisième cas on adopte la même géométrie que le premier cas et on change la chaleur spécifique thermoélectrique figure (4.6).

Sur la base de la structure adoptée du réseau, nous avons lancé la phase de formation à l'aide d'un algorithme d'entraînement Levenberg Marquardt. Les mesures d'impédance sont illustrées sur les figures (4.4),(4.5),(4.6). De plus, dans les mêmes conditions, nous tracerons également les valeurs de l'impédance obtenues par caractérisation par spectroscopie d'impédance de l'élément Bi_2Te_3 [69],[70] afin d'effectuer une comparaison entre les différents résultats et de mettre en évidence l'efficacité du modèle de réseau neuronal.

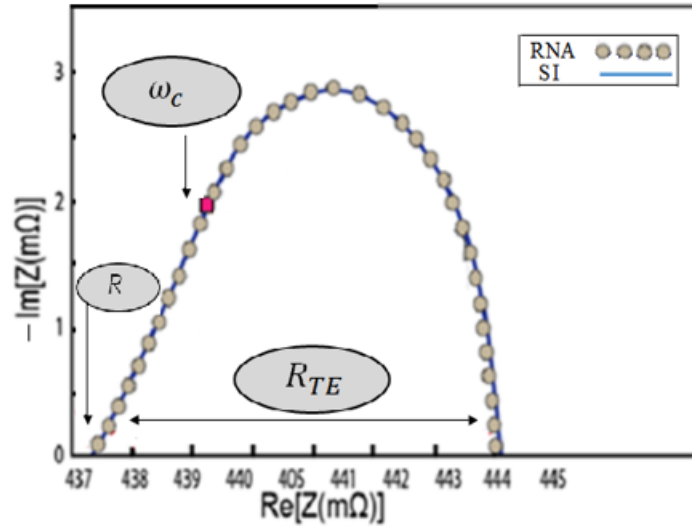


Figure 4.4 – Le spectre d'impédance de l'élément thermoélectrique Bi_2Te_3 . Les cercles se réfèrent aux spectres d'impédance dérivés de la réponse du réseau neuronal [98] et les lignes bleues font référence aux mesures par spectroscopie d'impédance effectuées par Jorge Garcia et al [70].

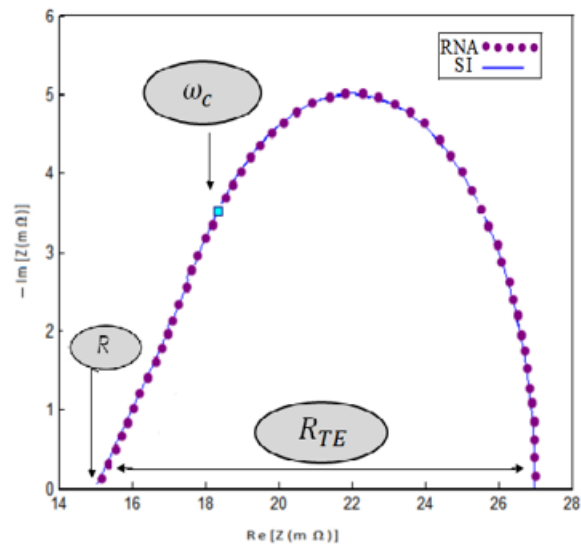


Figure 4.5 – Résultats simulés en termes de mesures d'impédance pour l'élément thermoélectrique Bi_2Te_3 . Les cercles se réfèrent aux spectres d'impédance dérivés de la réponse du réseau neuronal et les lignes bleues font référence aux mesures par spectroscopie d'impédance [69].

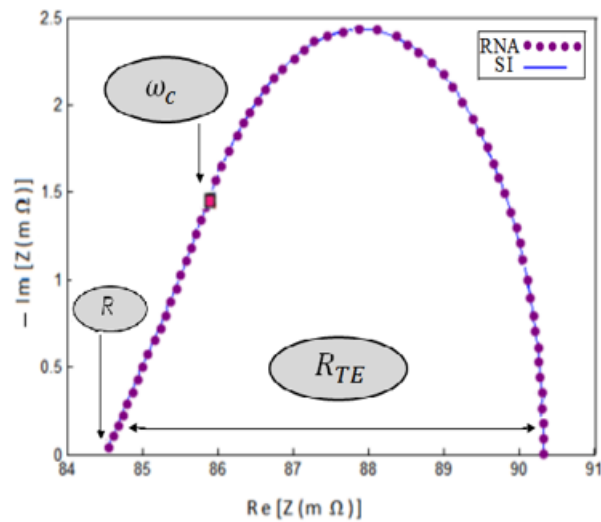


Figure 4.6 – Le spectre d'impédance de l'élément thermoélectrique Bi_2Te_3 . Les cercles se réfèrent aux spectres d'impédance dérivés de la réponse du réseau neuronal et les lignes bleues font référence aux mesures par spectroscopie d'impédance effectuées par Jorge Garcia et al [69].

Dans les figures (4.4),(4.5),(4.6) les courbes de Nyquist montrent que les composantes imaginaires (Im) et les composantes réelles (Re) de $Z^{IS}(j\omega)$ et $Z^{ANN}(j\omega)$ sont en très bonne adéquation, le modèle neural choisi a permis de reproduire les valeurs de l'impédance avec un excellent accord de ceux déduits par caractérisation par spectroscopie d'impédance. Ces résultats montrent l'efficacité de notre modèle de réseaux de neurones artificiels proposé pour dériver les mesures d'impédance du Bi_2Te_3 . Ces résultats mettent en évidence le bon choix de l'approche neuronale pour la description des matériaux thermoélectriques.

Pour discuter du comportement thermoélectrique à partir des graphes de l'impédance dans les figures (4.4),(4.5),(4.6), nous soutenons le fait que le comportement thermoélectrique est lié à la présence simultanée de courant électrique et de chaleur dans un métal ou un semi-conducteur sous l'effet d'un gradient de température. Le matériau est porteur de charges électriques (électrons, ions, trous), le flux de ces particules chargées et de la chaleur est de type diffusif. En analysant ces figures on distingue deux régimes de comportements différents séparés autour de la fréquence angulaire caractéristique ω_c :

Lorsque $\omega \gg \omega_c$ l'impédance a l'allure d'une ligne (pente = 1) décrite par l'expression théorique

$$Z = R + R_{TE} \left(\frac{j\omega}{\omega_{TE}} \right)^{-0,5} . \quad (4.2)$$

Sous ce régime, la variation de la température aux jonctions commence à s'établir, les deux ondes thermiques produites aux jonctions par la production et l'absorption de la chaleur induite par l'effet Peltier se propagent à l'intérieur du thermoélément sans qu'elles s'influencent, les ondes n'ont pas assez de temps pour pénétrer et atteindre la demi-longueur du thermoélément, d'après l'expression de la diffusivité thermique $\alpha_{TE} = \frac{\lambda_{TE}}{\rho C_p}$, l'absorption ou la libération de la chaleur à travers le matériau est exprimée par le produit de la masse volumique ρ par la chaleur spécifique C_p et la conduction thermique est exprimée à travers la conductivité thermique λ_{TE} et le gradient de température. C'est un régime de diffusion semi-

infini qui peut être déduit en résolvant la même équation de chaleur avec les conditions aux limites $\theta(x = \infty)$ et $\theta(x = 0)$:

$$\begin{cases} \theta(\infty, j\omega) = 0 \\ \theta(0, j\omega) = - \left[\left(\frac{ST_i i_0 L_H}{\lambda_{TE} A} \right) \right] \left(\frac{j\omega}{w_{TE}} \right)^{-1} \end{cases} \quad (4.3)$$

On obtient alors l'expression de l'onde thermique

$$\theta(x, j\omega) = \frac{ST_i i_0 L_H}{\lambda_{TE} A} \left(\frac{j\omega}{\omega_{TE}} \right)^{-0.5} \left\{ \sinh \left[\frac{x}{L_H} \left(\frac{j\omega}{\omega_{TE}} \right)^{0.5} \right] - \cosh \left[\frac{x}{L_H} \left(\frac{j\omega}{\omega_{TE}} \right)^{0.5} \right] \tanh \left[\left(\frac{j\omega}{\omega_{TE}} \right)^{0.5} \right] \right\} \quad (4.4)$$

Lorsque $\omega \ll \omega_c$ la réponse d'impédance se transforme en demi-cercle (figures (4.4),(4.5),(4.6) correspondant à l'équation

$$Z^{-1} = \frac{1}{R_{TE}} + \frac{1}{3} \frac{j\omega}{R_{TE} \omega_{TE}} \quad (4.5)$$

Dans ce cas les ondes thermiques ont suffisamment de temps pour pénétrer jusqu'à $L_H = \frac{L}{2}$, les amplitudes maximales se produisent toujours aux jonctions et décroissent en pénétrant vers L_H et la réponse d'impédance est régie par leur interaction. À mi-longueur, l'apport de chaleur du côté droit est égal à l'extraction de chaleur de la gauche et les ondes thermiques ne peuvent pas pénétrer plus loin, atteignant progressivement le régime de courant continu. On en conclut que, dans un thermoélément seul, les ondes thermiques diffusent semi-infiniment des jonctions vers la demi-longueur à hautes fréquences. Lorsque la fréquence est réduite, les ondes thermiques peuvent pénétrer d'avantage et finalement atteindre la demi-longueur où elles commencent à s'annuler et toute pénétration ultérieure est bloquée[70].

Sur la base de ces résultats certains paramètres peuvent être extraits, ce sont des connaissances qui fourniront un point de départ pour une bonne compréhension des résultats d'impédance dans les éléments thermoélectriques, permettant de mieux comprendre le phénomène physique et donnent une image complète des

dispositifs dans des conditions réelles de travail. En effet, pour un thermoélément (avec les mêmes hypothèses) mais en considérant maintenant cette fois-ci l'effet Joule, la solution de l'équation d'équilibre thermique en régime permanent du côté froid est donnée par

$$0 = \frac{\pi I}{A} + \frac{\lambda_{TE}}{L} (T_H - T_C) + \frac{1}{2} \frac{I^2 R}{A} \quad (4.6)$$

Ou encore

$$(T_H - T_C) = \frac{-L}{\lambda_{TE} A} \left(\pi I + \frac{1}{2} I^2 R \right) \quad (4.7)$$

Alors

$$S(T_C - T_H) = \frac{SL}{\lambda_{TE} A} \left(\pi I + \frac{1}{2} I^2 R \right) \quad (4.8)$$

La résistance thermoélectrique est le rapport de la tension Seebeck sur le courant

$$R_{TE} = \frac{V_S}{I} \quad (4.9)$$

Et comme

$$V_S = S(T_C - T_H) \quad (4.10)$$

On en déduit donc

$$R_{TE} = \frac{SL}{\lambda_{TE} A} \left(ST + \frac{1}{2} IR \right) \quad (4.11)$$

Lorsque l'effet joule est négligé :

$$R_{TE} = \frac{s^2 T_i L}{\lambda_{TE} A} \quad (4.12)$$

R_{TE} comporte les pertes du système introduites par les effets thermoélectriques, en effet lorsqu'on applique une différence de potentiel à $t=0$, une différence de température apparaît à l'équilibre due à l'effet Peltier, et les porteurs de charges se redistribuent en générant la tension Seebeck. Le courant Seebeck I_s est dans la direction opposée de I comme conséquence de l'apparition du voltage Seebeck on

peut définir la capacité thermique par

$$C_{TE} = I \frac{dt}{dv_s} = \frac{1}{s} \left(\frac{d(T_H - T_C)}{dt} \right)^{-1} \quad (4.13)$$

en utilisant la relation

$$dV_s = Sd(T_H - T_C) \quad (4.14)$$

Et en déduisant la conductivité thermoélectrique λ_{TE} de la relation de conduction de la chaleur et en introduisant l'expression de la chaleur spécifique C_p

$$\begin{cases} \frac{dQ}{dt} = -\lambda_{TE} \cdot A \frac{dT}{dx} \\ C_p = \frac{1}{\rho AL} \cdot \frac{dQ}{dT} \end{cases} \Rightarrow dt = \frac{-L}{\alpha} dx \quad (4.15)$$

On remplace cette dernière expression dans la relation (4.13)

$$C_{TE} = \frac{-IL}{S\alpha} \left(\frac{d(T_H - T_C)}{dx} \right)^{-1} \quad (4.16)$$

Et comme

$$d(T_H - T_C) = -2 dT_c \quad (4.17)$$

et en utilisant le bilan thermique du coté froid

$$\left(\frac{dT}{dx} \right)_L^{-1} = \frac{\lambda_{TE} A}{I\pi} \quad (4.18)$$

On obtient alors

$$C_{TE} = \frac{1}{R_{TE}\omega_{TE}} \quad (4.19)$$

Cette capacité, est déterminée par le processus thermique, elle contrôle la réponse imaginaire de l'admittance aux basses fréquences, ce qui induit que la tension Seebeck est $C_{TE} = \frac{Idt}{dv_s}$. On peut voir qu'à stable état ($\omega = 0$), l'impédance est $Z = R_{TE}$ et donc la distance entre les intersections avec l'axe réel fournit directement R_{TE} (figures (4.4),(4.5),(4.6)). Ce qui donne des informations sur la

séparation des charges et sur le courant de Seebeck qui apparait dans le système. Ainsi les nouveaux concepts de capacité thermique et de résistance thermique sont introduits avec la fréquence angulaire caractéristique. De plus, une constante de temps est déterminée par le produit des deux paramètres qui est liée au temps que prend le flux de chaleur pour diffuser dans le thermoélément.

$$\tau = C_{TE} \cdot R_{TE} = \frac{s^2 T_i L}{\lambda_{TE} A} \cdot \frac{\rho C_p A L}{2 S^2 T} = \frac{\left(\frac{L}{2}\right)^2}{\alpha}$$

Une fois les spectres d'impédance définis, on peut alors extraire les valeurs du coefficient Seebeck S , de la conductivité thermique λ_{TE} , de la résistance thermique R_{TE} , de la fréquence caractéristique de diffusion thermique ω et de la diffusivité thermique α_{TE} , ces résultats sont illustrés dans les tableaux (4.1)(4.2)(4.3). Il est évident d'après les figures (4.4), (4.5) et (4.6) que la réponse du réseau a bien défini les quantités liées aux propriétés thermoélectriques du Bi_2Te_3 . Les résultats du coefficient de Seebeck, de la conductivité thermique, de la résistance thermique et de la diffusivité thermique mettent en évidence l'étroite égalité entre les prévisions du réseau et les données par spectroscopie d'impédance [98, 69, 70] .

	λ_{TE} (w/M.k)	α_{TE} (cm^2/s)	C_{pTE} ($J/g.k$)	ω_{TE} (rad/s)	C_{TE} (F)	R_{TE} ($m\Omega$)	S ($\mu V/K$)	R ($m\Omega$)
Caractérisation par IS	1.27	0.013	0.13	2.01	86.21	5.83	175	84.4
Prédiction par RNA	1.23	0.014	0.12	1.98	86.95	5.81	171	84.1

Table 4.1 – Les valeurs des différents paramètres thermoélectriques de l'élément Bi_2Te_3 caractérisé par une surface de $1.4 \times 1.4 \text{ mm}^2$ et 1.6 mm de longueur.

	λ_{TE} (w/M.k)	α_{TE} (cm^2/s)	C_{pTE} (J/g.k)	ω_{TE} (rad/s)	C_{TE} (F)	R_{TE} ($m\Omega$)	S ($\mu V/K$)	R ($m\Omega$)
Caractérisation par IS	1.37	0.011	0.45	1.78	84.1	6.68	195	437
Prédiction par RNA	1.33	0.011	0.16	1.74	85.7	6.70	191	437

Table 4.2 – Les valeurs des différents paramètres thermoélectrique de l'élément Bi_2Te_3 caractérisé par une surface de $1.4 \times 1.4 mm^2$ et 1.6 mm de longueur.

	λ_{TE} (w/M.k)	α_{TE} (cm^2/s)	C_{pTE} (J/g.k)	ω_{TE} (rad/s)	C_{TE} (F)	R_{TE} ($m\Omega$)	S ($\mu V/K$)	R ($m\Omega$)
Caractérisation par IS	1.5	0.011	0.18	2	41.6	12	200	15
Prédiction par RNA	1.5	0.010	0.19	1.89	45.2	12	198	15

Table 4.3 – Les valeurs des différents paramètres thermoélectriques de l'élément Bi_2Te_3 caractérisé par une surface de $1mm^2$, et 1.5mm de longueur.

Validation expérimentale

Dans le but de valider expérimentalement les résultats de simulation, des mesures d'impédance ont été effectuées sur le thermoélément Bi_2Te_3 de type p (d'une section transversale $A = 1,4 \times 1,4mm^2$ et une longueur $L = 1,6mm$: géométrie choisie par Gracia et al.) par un dispositif expérimental décrit dans les références [69],[70]. Sur les figures (4.7) et (4.8) sont donnés les réponses expérimentales d'impédance.

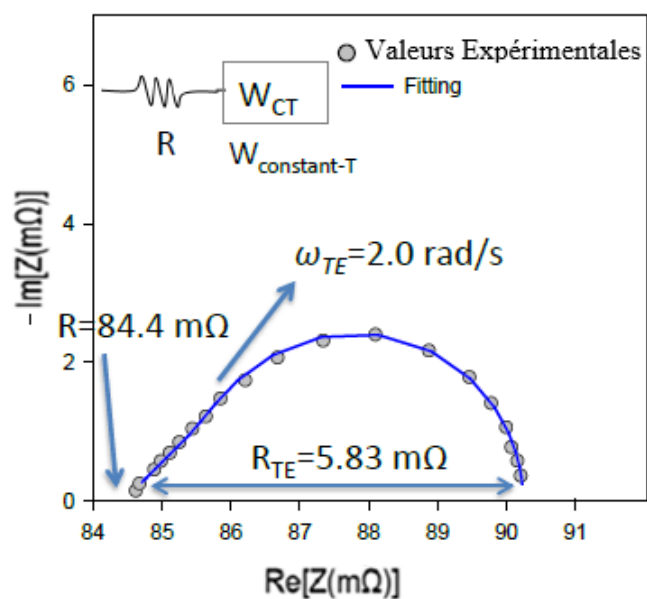


Figure 4.7 – Réponse expérimentale d'impédance (cercles) du Bi_2Te_3 de type p [69].

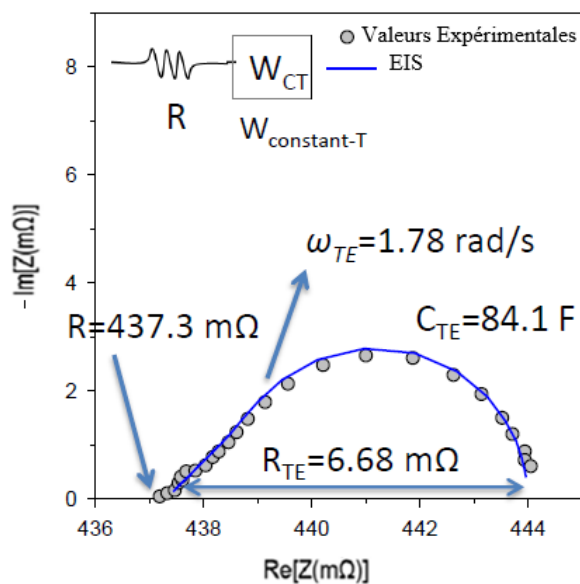


Figure 4.8 – Réponse expérimentale d'impédance (cercles) du Bi_2Te_3 de type p [70].

Il apparait clairement que les courbes expérimentales figure (4.7) et (4.8) montrent une similitude des formes avec les courbes obtenues dans les figures (4.6) et (4.4) qui contiennent les résultats de simulation par réseau de neurones artificiels.

4.3 Validation du modèle via la réponse du réseau

Les effets thermoélectriques étant présents dans tous les matériaux, la caractérisation des matériaux dont les propriétés thermoélectriques sont faibles constituerait un outil d'analyse pertinent. C'est le cas par exemple des métaux qui présentent une grande conductivité thermique et un faible coefficient Seebeck. En particulier les jonctions métalliques constituant les thermocouples présentent des propriétés intéressantes en tant que systèmes modèles pour la mesure. Ce sont des systèmes simples dont les propriétés sont bien connues. De plus, le rapport surface sur section des fils composant un thermocouple est bien plus grand comparativement à un système thermoélectrique classique, ce qui va induire une contribution beaucoup plus importante des pertes par convection. Quelques métaux ont donc été retenus pour entraîner notre modèle.

L'entraînement du modèle neuronal a été utilisé en collectant les base de données dans le cadre de l'impédance. Les collectes de données ont été extraites en définissant les valeurs empiriques dans les limites de la plage de chaque paramètre prévu. Pour l'étape de simulation, nous révélons principalement la capacité du modèle neuronal à prédire le coefficient Seebeck S_{TE} , la conductivité thermique λ_{TE} , la diffusivité thermique α_{TE} et le coefficient Peltier π de certains éléments. Nous avons testé la prédiction de notre modèle sur la base de certains métaux tels que l'argent, le cuivre, le cadmium, le thulium et le zinc. Dans le tableau (4.3), les prévisions globales du réseau sont comparées aux valeurs expérimentales relevées des littératures [99] à [102].

		Argent (Ag)	Cuivre (Cu)	Cadmium (Cd)	Thulium (Tm)	Zinc (Zn)
$S_{TE}(\mu V/K)$	RNA	1.57	1.72	2.29	1.8	1.7
	EXP	1.51	1.94	2.55	1.9	2.4
$\lambda_{TE}(W/m.K)$	RNA	431	325	74	16	107
	EXP	419	317	92	17	112
$\alpha_{TE}(Cm^2/s)(10^{-7})$	RNA	2.19	4.7	10.59	38.5	10.57
	EXP	2.42	3.1	8.69	35.7	10.14
$\pi(\mu V.s/K)$	RNA	471	532	747	556	691
	EXP	453	517	765	570	720

Table 4.4 – Les valeurs du coefficient de Seebeck S , de la conductivité thermique de la diffusivité thermique et du coefficient de Peltier de quelques éléments.

Les valeurs prédites par le reseau neronal sont les sorties du network, les valeurs dans la littérature sont obtenus des mesures expérimentales dans les références [99] à [102], toutes les valeurs sont obtenues à 300^0K . Comme le montre le tableau 1, le modèle de réseau de neurone a fourni les quantités de sortie des propriétés thermoélectriques à 300^0K . Nous avons illustré les sorties du réseau liées à certains éléments de test tels que Ag, Cu, Cd, Tm et Zn. Selon les sorties calculées, on peut affirmer que les prédictions du réseau sont en accord avec les calculs expérimentaux. En d'autres termes, on peut affirmer que l'architecture adoptée, dans le cadre de l'impédance, est capable de calculer les propriétés thermoélectriques des éléments avec une précision adéquate. Il est également important de préciser que le modèle ne peut prédire que des valeurs positives des coefficients Seebeck, car nous n'avons adopté que des valeurs positives lors de la formation en état de marche et ceci pour des raisons de simplicité. Nous définissons également les performances de la topologie du réseau neuronal pour l'entraînement de l'ensemble de données en calculant l'erreur quadratique moyenne qui définit la réponse du réseau par rapport à l'ensemble de données d'apprentissage. La topologie de réseau con-

struite contient deux couches et les paramètres réseau sont initialisés et produits en utilisant l'algorithme de Nguyen et Widrow.

Il est clair sur la figure (4.9) que le réseau défini donne une réponse efficace

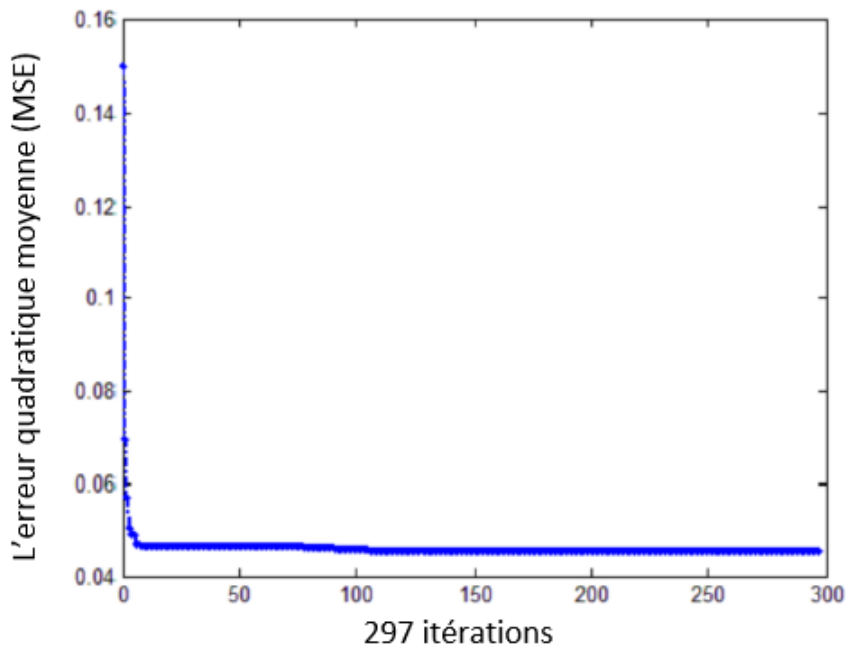


Figure 4.9 – L'erreur quadratique moyenne par rapport au nombre d'itérations du modèle neuronal.

pendant la phase d'apprentissage. En fait, les unités d'entrée-sortie entraînées de l'approche proposée montrent une architecture stable qui fournit un modèle efficace capable de prédire les différentes propriétés thermoélectriques. Comme le montre la figure (4.9), les valeurs entraînées données par l'approche sont en accord par rapport aux cibles souhaitées. Dans ce cas, à la fin du processus d'apprentissage, nous avons obtenu $MSE \rightarrow 10^{-2}$. Par conséquent, nous pouvons affirmer que la capacité du réseau neuronal à prédire les propriétés thermoélectriques est atteinte.

4.4 Conclusion

Dans ce travail, nous avons présenté un modèle de réseau de neurones artificiel qui peut être appliqué pour caractériser les matériaux thermoélectriques. L'architecture du réseau neuronal a été développée sur la base du critère de sélection des unités d'entrée-sortie. C'est une méthode basée sur une nouvelle approche de codage pour générer l'ensemble (entrée-sortie) lors du processus d'apprentissage. Cette façon d'intégrer les datas a permis de trouver les meilleurs paramètres optimaux qui correspondent au traitement de ces datas. La topologie du réseau, choisie pour moduler les quantités d'intérêt a intégré les données d'impédance comme unités d'entrée du réseau et les unités de sortie incluent le coefficient de Seebeck et la conductivité thermique. En raison de la convergence locale liée à chaque condition aux limites, un ensemble aléatoire de poids initiaux et de valeurs de biais a été développé. Les paramètres du réseau ont été mis à jour à l'aide de l'algorithme de Levenberg Marquardt et l'algorithme de rétropropagation, sur la base de l'approche adoptée, les performances du réseau sont évaluées en illustrant les tracés de Nyquist montrant les comportements d'impédance à différentes fréquences, où les propriétés thermoélectriques ont été déduites.

La comparaison entre la réponse expérimentale et celles obtenues par les valeurs issues des deux approches : caractérisation neuronale et caractérisation par spectroscopie d'impédance ont confirmé l'efficacité de cette technique d'évaluation et a donné des résultats tout à fait comparables, en effet, une bonne corrélation a été obtenue par les paramètres issus d'une approche neuronale par rapport à la réponse expérimentale. Les réseaux de neurone sont un outil puissant capable de résoudre une variété de tâches, l'approximation des fonctions, la reconnaissance des caractères et la modélisation des problèmes physiques. Leur puissance est utilisée pour développer des solutions alternatives de plus en plus efficaces. Notre travail a confirmé l'efficacité de cette technique d'évaluation et a donné des résultats tout à fait comparables.

Conclusion Générale

La modélisation mathématique est un outil indispensable à la représentation de systèmes pour maîtriser ou optimiser leur fonctionnement, ou comprendre leur évolution. Elle nécessite des mesures sur le système à modéliser et la connaissance des lois régissant son fonctionnement et qui a fait l'objet de ce travail de thèse.

L'objectif de notre contribution portait sur l'utilisation des Réseaux de Neurones Artificiels pour la caractérisation des propriétés thermoélectrique des éléments, à cause de la nature fortement non linéaire des phénomènes thermoélectriques, cette approche présente plusieurs avantages, d'une part la possibilité de prendre compte de plusieurs paramètres en même temps, d'autre part la possibilité d'étudier l'intégralité de ces propriétés : le coefficient Seebeck, la conductivité thermique, la résistivité électrique, mais également la diffusivité thermique, et les résistances thermiques d'interface, ces derniers ont été rarement étudiées, puisque elles sont difficiles à mesurer.

Dans un premier temps, (le premier et le deuxième chapitre), une description des réseaux connexionnistes a été donnée. Un aperçu sur l'apprentissage des différents modèles ainsi que les méthodes d'apprentissage qui s'appuient sur la rétro propagation de l'erreur ont été alors rapportés. Nous avons, par la suite, dégagé les propriétés spécifiques aux réseaux de neurones qui sont à l'origine de leur intérêt pratique toujours croissant. Cette partie s'est achevée par une revue générale des

domaines d'applications.

Le troisième chapitre a débuté par une introduction aux domaines pratiques de l'utilisation des éléments thermoélectriques comme générateurs électriques ou comme éléments de réfrigération en abordant leur principe de fonctionnement et en discutant de la place qu'a actuellement la thermoélectricité dans le secteur de l'énergie. Ensuite nous avons décrits les modèles d'impédance électrique requis pour interpréter correctement les résultats expérimentaux en résolvant l'équation de la chaleur dans le domaine fréquentiel dans des conditions adiabatiques et fonctionnement en mode Peltier d'un dispositif thermoélectrique, l'analyse dans le plan complexe de l'expression de l'impédance fournit l'élément de circuit équivalent requis pour interpréter les mesures de l'impédance. En utilisant l'approche de caractérisation par spectroscopie d'impédance toutes et les propriétés thermiques peuvent être extraites à une température donnée.

Dans le quatrième chapitre nous avons développé une approche neuronale capable de décrire certaines propriétés telles que le coefficient Seebeck, la conductivité thermique, la résistance thermique, la fréquence caractéristique de diffusion thermique et la diffusivité thermique, la topologie du réseau choisie pour moduler les quantités d'intérêt des éléments thermoélectriques, a intégré les données de l'impédance sous sa forme complexe comme unités d'entrée pour simuler les propriétés électrochimiques. Nous avons formé notre modèle du réseau de neurone artificiel pour saisir les paramètres du réseau qui définissent mieux le modèle optimal en utilisant la technique d'apprentissage supervisé et des algorithmes de second ordre pour mettre à jour les paramètres de pondération de notre modèle. Par conséquent, la capacité de prédire les propriétés thermiques est atteinte. Basé sur la représentation d'impédance des matériaux thermoélectriques déduite de la résolution de l'équation de la chaleur, nous visons à définir une réponse globale liée à l'ensemble de données d'entrée et aux paramètres de réseau. Cette étude a permis de démontrer l'intérêt des réseaux de neurones comme outil de modélisa-

tion.

L'objectif était de proposer une méthodologie pour développer un modèle à partir de réseaux de neurones selon deux axes : d'une part l'architecture est définie en fonction des constituants et des conditions opératoires sensibles (température, forme, ...) et d'autre part les valeurs des paramètres (poids) sont identifiées à partir d'un ensemble de données disponibles. Une étude comparative entre la réponse expérimentale et celles obtenues par les valeurs identifiées issues des deux approches est présentée : l'approche neuronale et l'approche par spectroscopie d'impédance. La comparaison a confirmé l'efficacité de notre technique d'évaluation et a donné des résultats tout à fait comparables avec parfois une réduction très appréciable du temps de calcul, en effet, une bonne corrélation a été obtenue par les paramètres issus de l'approche neuronale par rapport à la réponse expérimentale. Nous avons pu apprécier ainsi les avantages de celle-ci par rapport à la méthode d'optimisation classique qui sont :

- une bonne concordance a été constatée,
- un gain de temps calcul et une rapidité notable,
- il n'y a pas de risque que les variables soient déterminées en dehors de leurs domaines admissibles.

Notre travail a confirmé l'efficacité de cette technique d'évaluation et a donné des résultats tout à fait comparables. Les perspectives du présent travail de recherche sont nombreuses. On se propose d'étendre notre analyse pour tenir compte de l'effet de convection à la surface externe des plaques céramiques des modules thermoélectriques, qui est la partie du dispositif où la convection est la plus importante. Ceci est réalisé en résolvant l'équation de chaleur dans le domaine fréquentiel, y compris les pertes de chaleur par convection. On peut aussi appliquer notre approche sur les skutterudites thermoélectriques qui sont de bons éléments lorsque les températures sont élevées, encore, il pourrait être envisagé

de développer un modèle de réseau de neurones permettant de générer un modèle de comportement bien déterminé. Cela permet en fait de passer à des modèles plus réalistes et plus fins puisqu'il nous offre la possibilité de prendre compte de plusieurs paramètres en même temps.

A partir des résultats trouvés dans ce travail et la manipulation de cet outil tout au long de notre travail de thèse nous avons constaté que le véritable avantage de cette technique était dans leur faculté à représenter des rapports linéaires et non linéaires complexes, et dans sa capacité à apprendre ces rapports directement de la série de données modélisée. Cette propriété permet aux réseaux de neurones d'être un bon candidat pour prendre des entrées dans un nouvel ensemble et pour produire la sortie correspondante dans un temps de simulation très court et cela fait beaucoup de temps de calcul. Ainsi, les réseaux de neurones, dotés peu à peu de supports matériels de plus en plus adaptés et puissants, mieux étudiés et mieux connus, deviendront une véritable technologie nouvelle.

Bibliographie

- [1] D. Panazoli. Proposition de l'architecture Cortexionist pour l'intelligence comportementale de créatures artificielles. Thèse de Doctorat 2008. Université de Toulouse. France.
- [2] A. Belayadi, Généralisation de l'approximation de tâches par les réseaux de neurones à plusieurs variables en utilisant les algorithmes de second-ordre. Thèse de Doctorat 2018. USTHB. Algérie.
- [3] W. S. McCulloch, W. Pitts. A logical calculus of ideas immanent in nervous activity. Bulletin of mathematical biophysics, 5 (PP :115-133). MIT Press. 1943.
- [4] D.Hebb. The Organisation Of Behavior. Wiley. 1949.
- [5] F. Rosenblatt. The perceptron : probabilistic model for information storage and organization in the brain. Psychological Review 65 (PP :386-408) Perceptron Monocouche. 1958.
- [6] J. Hopfield. Neurons with graded response have collective computational properties like those of two-states neurons. Proceedings of the National Academy of Sciences 81.(PP :3088-3092). 1984.
- [7] B. Irie, S. Miyake. Capabilities of tree-layered perceptrons. Proceedings of the IEEE Second International Conférence on Neural Networks (San Diego), 1 (PP :641-647). 1988.
- [8] M. Parizeau, Réseaux de neurones. GIF-21140 et GIF-64326. Université de Laval. 2004.

- [9] L. Ait Gougam. Application des réseaux de neurones-ondelettes à l'approximation de tâches et à l'analyse de certaines oscillations localisées. Thèse de Doctorat 2009. USTHB. Algérie.
- [10] F.M. Hum, I. Kostanic. Principles of neurocomputing for science and engineering, 2^{ed}. McGraw Hill. 2001. New York
- [11] I. Rivals, L. Personnaz, G. Dreyfus, J.L. Ploix. Modélisation, classification et commande par réseaux de neurones : principes fondamentaux, méthodologie de conception et applications industrielles, Les réseaux de neurones pour la modélisation et la commande de processus, J.P. Corriou édition. La voisier Tec et Doc. 1995.
- [12] 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.(PP :229-237). IEEE Press. 1994.
- [13] C. Touzet. Les réseaux de neurones artificiels : Introduction au connexionnisme. Tech rep. École pour les études et la recherche en informatique et électronique. Nîmes. (1992).
- [14] A. Benveniste, Q. Zhang. Wavelet networks. IEEE Trans. Neural Networks. (PP :68-89). 1992.
- [15] S. Grossberg. Adaptive pattern classification and universal recoding, Parallel development and coding of neural feature detectors : Feedback, expectation, olfaction and illusions. Biological Cybernetics, 23. 1976.
- [16] T. Kohonen. Self-organization and Associative Memory. Springer-Verlag, Berlin. 1984.
- [17] K. Hornik, M. Stincombe, H. White. Multilayer feedforward networks are universal approximators, Neural Networks. 2. (PP :359-366). 1989.
- [18] A. Krizhevsky, G. Sutskever, E. Hinton. Image net classification with deep convolutional neural networks. Advances in neural information processing systems.(PP :1097-1105). 2012.

- [19] O. Nerrand, P. Roussel-Ragot, L. Personnaz, G. Dreyfus, Neural networks and non linear adaptive filtering : unifying concepts and new algorithms *Neural Computation*, 5. (PP : 165-199). 1993.
- [20] D.E. Rumelhart, G.E. Hinton, R.J. Williams. Learning representations by back-propagating errors. *Nature*, 323. (PP :533-536). 1986. London.
- [21] G.Y. le Cun. Modèles connexionnistes de l'apprentissage. Thèse de Doctorat. Université Pierre et Marie Curie, Paris. 1987. France.
- [22] K. Hornick , M.stinchcombe, H.White. Multilayer feedforward neural networks universal aproximators. *Neural Networks* 2, (PP :359-366). 1989.
- [23] K. Funakashi. On the approximate relization of continious mappings by neural networks. *Neural networks* 2,(PP :183-214). 1989.
- [24] P.D. Mandic, J.A Chambers. Recurrent neural networks for prediction Learning algorithms, architectures and stability. J wiley and Sons, LTD. 2001.
- [25] S. Hochreiter, J. Schmidhuber. Long short term memory. *Neural Computation*, 9.(PP :1735-1780). 1997.
- [26] J. Michael, D. Powell. Radial basis functions for multi-variable interpolation : A review IMA Conference on Algorithms for the approximation of functions and data RMCS Shrivenham.(PP :143-167). 1985. UK.
- [27] J. Park, IW. Sandberg. An approximation and radial-basis-function networks. *Neural Computation*, 5. (PP :305-316). 1993.
- [28] J. Moody, C. Darken. Fast learning in networks of locally-tuned processing units. *Neural Computation*. (PP : 281-294). 1989.
- [29] G.A. Carpenter, S. Grossberg. The Handbook of Brain Theory and,,Neural Networks, Second Edition, chap. Adaptive resonance theory, (PP : 87-90).MIT Press, Cambridge, MA, arbib edn, 2003.
- [30] R. Fletcher. Practical methods of optimization. John Wiley and Sons, New-York,1987.

- [31] R. Fletcher. A New approach to variable metric algorithms. *Computer Journal*, 13, (PP :317-322), 1970.
- [32] A.V. Fiacco, G.P. McCormick. *Nonlinear programming : sequential unconstrained minimization techniques*. John Wiley, New-York, 1968.
- [33] M. Minoux. *Programmation mathématique : théorie et algorithmes*. Dunod, 1983.
- [34] A. Björck. *Numerical methods for least squares problems*. SIAM, Philadelphia, 1996.
- [35] J.B. Rosen. The gradient projection method for nonlinear programming, part 1 : linear constraints. *SIAM Journal*, 8. (PP : 181-217). 1960.
- [36] D.E. Rumelhart, G.E. Hinton, R.J. Williams. Learning representations by back-propagating errors. *Cognitive modeling* 5(3). 1988.
- [37] R. Battiti. First and second order methods for learning between steepest descent methods and Newtons method. *Neural Computation*, 4, (PP : 141-166). 1992.
- [38] H.W. Kuhn, A.W. Tucker. *Linear inequalities and related systems*. Princeton University Press, 1956.
- [39] H.W. Kuhn, A.W. Tucker. Nonlinear programming. In *Proc. of the 2nd Berkeley Symp on mathematical statistics and probability*. University of California, Press Berkeley. (PP : 481-492). 1951.
- [40] J.E. Dennis, J.J. More. Quasi-Newton methods, motivation and theory. *SIAM Review*, 19. (PP : 46-89). 1977.
- [41] J.J. Moré. The Levenberg-Marquardt Algorithm : Implementation and Theory. In *Numerical Analysis, Lecture Notes in Mathematics*, 630. Springer Verlag. (PP : 105-116). 1977.
- [42] E. Walter, L. Pronzato. *Identification de modèles paramétriques à partir de données expérimentales*. Editions Masson, Paris. 1994.

- [43] M.T. Hagan, M. B. Menhaj. IEEE Trans. Neural Network, 5. (PP :989-993). 1994.
- [44] S. Samarasinghe, Neural Networks for applied Sciences and Engineering From Fundamentals to Complex Patten Recognition. Taylor and Francis Group, LLC. 2007.
- [45] B. Kroise, P. Van Der Smaget. An introduction to neural networks. 8^{ed}. The University of Amsterdam. 1996.
- [46] G. Cybenko, Mathematics of Control, Signals and Systems 2, 303. 1989.
- [47] G. Dreyfus, J.M. Martinez, M. Samuelides, M.B. Gordon, F. Badran, S. Thiria, L. Hérault . Réseaux de neurones, méthodologie et applications Editions Eyrolles, Paris. 2002.
- [48] K. Hornick, M. Stinchombe, H. White, P. Auer. Degree of approximation results for feedforward networks approximating unknown mappings and their derivatives, Neural Computation, 6. (PP : 1262-1275). 1994.
- [49] A. Belayadi, A. Mougari, M. Zabat. Modeling of electrochemical properties of potential-induced defects in butane-thiol sams by using artificial neural network and impedance spectroscopy data, J. Solid State Electrochem (PP :195-204). 2018.
- [50] S. Hadjimi, A. Belayadi, M. Zabat, A. Mougari, M.W. Khemici, Impedance spectroscopy and neurocomputing approaches to investigate the enhanced electrical blocking properties of $\text{CH}_3(\text{CH}_2)_{n-1}\text{SH}$ thin monolayers electrodeposited on gold electrode, J. Electron. Mater. 48, (PP :3908-3918). 2019.
- [51] A. Mougari, M. Zabat, S. Boudjadar, Monitored deformation by potential-induced defects in $\text{CH}_3(\text{CH}_2)_{n-1}\text{SH}$ self-assembled monolayers on gold : IS study of chain length effect on subdiffusion and sams heterogeneity. Surface Review and Letters. 26(10). 19500672019. 2019.
- [52] X.C. Xuan, Optimum design of a thermoelectric device. Semiconductor science and technology, 17(2) : p. 114. 2002.

- [53] F. Ioffe. Physics of Semiconductors. Infosearch Ltd. London. 1957.
- [54] S.B. Riffat, X. Ma, Thermoelectrics : a review of present and potential applications, Applied Thermal Engineering. 23. (PP. 913-935). 2003.
- [55] C.B. Vining, Semiconductors are cool, Nature. 413, (PP : 577-578). 2001.
- [56] A. Melnikov, V.G. Kostishyn, V.V. Alenkov, Dimensionless model of a thermo-electric cooling device operating at real heat transfer conditions : maximum cooling capacity mode, J. Electron. Mater. 46 (PP :2737-2745). 2017.
- [57] A. Sherchenkov, Y.I. Shtern, R. Mironov, M.Y. Shtern , M.S. Rogachev, Current state of thermo-electric material science and the search for new effective materials, Nanotechnol. Russ. 10 (PP :827-840). 2015.
- [58] L. Ding, W. Ruochen, Y. Wei, Z. Weiqi, Performance evaluation of a novel thermo-electric module with BiSbTeSe-based material, Appl. Energy 238 (PP :1299-1311). 2019.
- [59] T.J. Seebeck. Über den magnetismus der galvenische kette, abh.k. Akad. Wiss. Berlin, 289. 1821.
- [60] G. Savelli. Etude et développement de composants thermoélectriques à base de couches minces. Thèse de Doctorat. Université Joseph-Fourier-Grenoble I. France. 2007
- [61] J.C.A. Peltier. Nouvelles expériences sur la caloricité des courants électrique. Ann,Chim. Phys, 56 (PP :371-386). 1834.
- [62] W. Thomson. On a mechanical theory of thermo-electric currents. Proceedings of the Royal society of Edinburgh, 3 (PP :91-98). 1857.
- [63] Y. Bouyrie, Identification, synthèse et caractérisation de phases tétraédrites pour la conversion d'énergie par effets thermoélectriques. Thèse de Doctorat. Université de Lorraine. 2015. France.

- [64] G. Min, D. Rowe, Recent concepts in thermoelectric power generation in Thermoelectrics. Proceedings ICT'02 Twenty-First International Conference, (PP :365-374). 2002.
- [65] E. Hatzikraniotis, K.T. Zorbas, I. Samaras, T. Kyratsi, K.M. Paraskevopoulos, Efficiency study of a commercial thermo-electric power generator (teg) under thermal cycling, J. Electron. Mater, 39. 2010.
- [66] H.J. Goldsmid. Bismuth Telluride and Its Alloys as Materials for Thermoelectric Generation Materials, 7(4) (PP :2577-2592). 2014.
- [67] J. R. Sootsman, D.Y. Chung, M.G. Kanatzidis. New and Old Concepts in Thermoelectric Materials. Chemistry. 48,(PP :8616-8639). 2009.
- [68] E. Thiebaut. Spectroscopie d'impédance non linéaire appliquée aux matériaux et systèmes thermoélectriques. Thèse de Doctorat. Université Paris-Saclay. France. 2019.
- [69] J. García-Canadas, G. Min, Impedance spectroscopy models for the complete characterization of thermo-electric materials, J. Appl. Phys. 116, 174510. 2014.
- [70] J. García-Canadas, G. Min, Thermal dynamics of thermo-electric phenomena from frequency resolved methods, AIP Advances, J. Appl. Phys 6, 035008. 2016.
- [71] E. Barsoukov, S.H. Ryu, H. Lee. A novel impedance spectrometer based on carrier function Laplace-transform of the response to arbitrary excitation. J of Electroanal Chemi 536,(PP :109-122). 2002.
- [72] V. Alar, I. Zmak, B. Runje, A. Horvati. Development of models for prediction of corrosion and pitting potential on AISI 304 stainless steel in different environmental conditions. Int. J. Electrochem. Sci, 11. (PP : 7674-7687). 2016.
- [73] J. Youshia, M.E. Ali, A. Lamprecht Artificial neural network based particle size prediction of polymeric nanoparticles. Eur. J. Biopharm, 119.(PP :333-342). 2017.

- [74] M.A.Z. Raja, Solution of the one-dimensional Bratu equation arising in the fuel ignition model using ANN optimised with PSO and SQP. *Connect Science*, 26(3). (PP :195-214). 2014.
- [75] G.Z. Quan, T. Wang, Y.L. Li , Z.Y. Zhan, Y.F. Xia. Artificial neural network modelling to evaluate the dynamic flow stress of 7050 aluminium alloy. *Journal of Materials Engineering and Performance*, 25(2). (PP :553-564), 2016.
- [76] H. Li, Z. Zhang, Z. Liu. Application of artificial neural networks for catalysis. *Catalysts* 7(10). 2017.
- [77] D.K. Kasdekar, V. Parashar, C. Arya . Artificial Neural Network Models for the prediction of MRR in electro-chemical machining. *Mater Today Proc*, 5(1). (PP :772-779). 2018.
- [78] A. Lazrak, F. Boudehenn, S. Bonnot, F. Fraisse, A. Leconte, P. Papillon, B. Souyri, Development of a dynamic artificial neural network model of an absorption chiller and its experimental validation. *Renew Energy*, 86. (PP : 1009-1022). 2016.
- [79] P.C. Verpoort, P. MacDonald, G.J. Conduit, Materials data validation and imputation with an artificial neural network. *Comput. Mater. Science*, 147. (PP :176-185). 2018.
- [80] M.A.Z. Raja, R. Samar, Numerical treatment for nonlinear MHD Jeffery-Hamel problem using neural networks optimized with interior point algorithm. *Neurocomputing*, 124. (PP :178-193). 2014.
- [81] M.A.Z. Raja, J.A Khan, N.I Chaudhary, E. Shivanian. Reliable numerical treatment of nonlinear singular FlierlPetviashvili equations for unbounded domain using ANN, GAs and SQP. *Appl Soft Comput*, 38 (PP :617-636). 2015.
- [82] M.A.Z. Raja, U. Farooq, N.I. Chaudhary, A.M. Wazwaz. Stochastic numerical solver for nanofluidic problems containing multi-walled carbon nanotubes. *Appl Soft Comput*, 3. (PP :561-586). 2016.

- [83] M.A.Z. Raja, R. Samar. Bio-inspired computing platform for reliable solution of Bratu-type equations arising in the modelling of electrically conducting solids. *Appl. Math. Model*, 40. (PP :5964-5977). 2016.
- [84] M.A.Z. Raja, J.A. Khan, T. Haroon. Stochastic numerical treatment for thin film flow of third grade fluid using unsupervised neural networks. *J. Taiwan Inst Chem Eng*, 48.(PP :2639). 2015.
- [85] M.A.Z. Raja. Stochastic numerical techniques for solving Troesch's problem. *Inf. Sci*, 279. (PP :860-873). 2014.
- [86] D. Colorado-Garrido, D.M. Ortega-Toledo, J.A. Hernández, J.G. González-Rodríguez, J. Uruchurtu. Neural networks for Nyquist plots prediction during corrosion inhibition of a pipeline steel. *J Solid. State. Electro-chem*, 13(11),(PP :1715-1722). 2009.
- [87] M. Nikoo, L. Sadowski, Prediction of the corrosion current density in reinforced concrete using a self-organizing featuremap. *Coatings*, 7(10). 2017.
- [88] A. Belayadi, B. Bourahla, F. Mekideche-Chafa Neurocomputing techniques to predict the 2d structures by using lattice dynamics of surfaces. *Acta Phys Pol A*, 132(4).(PP :1314-1319). 2017.
- [89] A. Belayadi, L. Ait-gougam, Bourahla B, F. Mekideche-Chafa. An approach based on neural computation to simulate the transition metals using tight binding measurements. *Turk. J. Phys*, 40. (PP :231-243). 2016.
- [90] A. Belayadi, L. Ait-Gougam, F. Mekideche-Chafa. Automatic pattern recognition with wavelet neural network. 2015.
- [91] A. Belayadi, Computing with an artificial neural network to enhance information processing : using a new methodology of feeding the training input-output mapping in, Nova Science Publishers. (PP : 193-223). 2018. USA.
- [92] A. Belayadi, B. Bourahla, A. Mougari, F. Mekideche-Chafa, Neurocomputing approach to predict 2D ferromagnetic lattices by spin dynamics. *Inte. J. of Model, Simul, Sci, Comput*, 11(5). 2020.

- [93] Y. Du, J. Xu, P.Eklund, B. Paulb, Flexible thermo-electric materials and devices, *Appl. Mater. Today*, 12. (PP :366-388). 2018.
- [94] L.C. Ding, A. Akbarzadeh, L. Tan. A review of power generation with thermo-electric system and its alternative with solar ponds, *Renew. Sustain. Energy Rev*, 81. (PP :799-812). 2018.
- [95] A.A. Melnikov, V.G. Kostishin, V.V. Alenkov, Dimensionless model of a thermo-electric cooling device operating at real heat transfer conditions : maximum cooling capacity mode, *J. Electron. Mater*, 46. (PP :2737-2745). 2017.
- [96] L. Ding, W. Ruochen, Y. Wei, Z. Weiqi, Performance evaluation of a novel thermo-electric module with BiSbTeSe-based material, *Appl. Energy*, 238. (PP :1299-1311). 2019.
- [97] B. Beltran-Pitarch, J. Prado-Gonjal, A. V. Powell, F. Martínez-Julian, J. García-Canadas. Characterization of Thermoelectric Materials by Impedance Spectroscopy, *J. Phys. Chem.* 123, (PP :12608-12613). (2019).
- [98] S. Zaoui, A. Belayadi, M. Zabat, A. Mougari, F. Mekideche-Chafa, An implementation of impedance spectroscopy approach to predict thermoelectronic properties of elements using a connectionist model of artificial neural network, *Physica B : Physics of Condensed Matter*, 580. 411735. USTHB. Algérie. 2020.
- [99] D.M. Rowe, *CRC Handbook of Thermo-Electric*, 2^{ed}, CRC-Press. 1995.
- [100] M. Yoshinao, S. Kimura, K. Wng, Enhancement of thermo-electric figure of merit through nanostructural control on intermetallic semiconductors toward high-temperature applications, *Nanomater. Res. appl.* 2006
- [101] L. Ho-Sung, *Thermo-electrics : Design and Materials*, John Wiley and Sons. Ltd. 2017.
- [102] A. Da-Rosa, *Fundamentals of Renewable Energy Processes*, 3^{ed}. Elsevier. 2013.