

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE  
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE  
SCIENTIFIQUE

*Université de Mohamed El-Bachir El-Ibrahimi - Bordj Bou Arreridj*

*Faculté des Sciences et de la technologie*

*Département d'Electronique*

# *Mémoire*

*Présenté pour obtenir*

**LE DIPLOME DE MASTER**

**FILIERE : Télécommunications**

**Spécialité : Systèmes des Télécommunications**

Par

**Brahimi Romayssa**

**Bechami NourELHouda**

*Intitulé*

*Diagnostic de la maladie de Parkinson basé sur l'analyse de la voix et  
l'application de l'apprentissage profond*

*Soutenu le 1/07/2025*

*Devant le Jury composé de :*

<i>Nom &amp; Prénom</i>	<i>Grade</i>	<i>Qualité</i>	<i>Etablissement</i>
<b>Mme.LAKHLEF NOURA</b>	<b>MCB</b>	<b>Président</b>	<b>Univ-BBA</b>
<b>Mme.ADOUI IBTISSEM</b>	<b>MCA</b>	<b>Encadreur</b>	<b>Univ-BBA</b>
<b>Mr. HACINE GHARBI ABDENOUR</b>	<b>MCA</b>	<b>Co-Encadreur</b>	<b>Univ-BBA</b>
<b>Mr. BOUDCHICHE DJAMELELDINE</b>	<b>MCA</b>	<b>Examineur</b>	<b>Univ-BBA</b>

*Année Universitaire 2024/2025*

## **Remerciement**

*Nous tenons à exprimer nos remerciements les plus sincères à Dieu, le Tout-Puissant, pour Ses innombrables bienfaits et pour nous avoir accordé la force, la patience et la persévérance nécessaires à l'accomplissement de ce travail*

*Nous adressons notre profonde gratitude et nos sincères remerciements à **Mr. HACINE GHARBI ABDENOUR** pour son soutien constant et ses conseils avisés qui ont grandement contribué au développement de ce travail. Nous avons beaucoup bénéficié de son expertise scientifique, de son souci de la rigueur et de la méthode, et ses orientations ont été présentes tout au long des différentes étapes de ce projet, ce qui a enrichi notre parcours de recherche et a aidé à surmonter les difficultés rencontrées.*

*Nous exprimons notre gratitude à **Mme. ADOUI IBTISSEM** pour ses remarques précieuses et sa collaboration qui ont contribué à améliorer la qualité de ce travail. Nous apprécions son soutien et sa contribution durant la période d'élaboration de ce projet.*

*Nous adressons également nos vifs remerciements aux éminents membres du jury pour avoir accepté d'évaluer ce travail et pour l'intérêt qu'ils lui ont accordé.*

*Enfin, nous tenons à remercier du fond du cœur toutes les personnes qui nous ont apporté leur soutien, ne serait-ce qu'à travers une parole encourageante, et qui nous ont aidés à garder la motivation nécessaire pour atteindre nos objectifs.*

## إهداء

إلهي لا يطيب الليل إلا بذكرك، ولا يهنأ النهار إلا بطاعتك، ولا تطيب الأخوة إلا بعفوك، ولا تكتمل الجنة إلا برويتك.

إلى من بلّغ الأمانة، ونصح الأمة، إلى نبيّ الرحمة ونور العالمين، محمد صلى الله عليه وسلم، أزمى الصلاة والسلام.

إلى جداتي الغاليات، التي كانتا نورًا وبركة لبيتنا، وفخرًا لي أنني حفيدتهما، رحمك الله يا "جدتي فريحة"، ورحمك الله يا "جدتي ساسية"، وأسكنكما فسيح جناته.

إلى والدي العزيز، الرجل الأول في حياتي، وسندي ومن منحني اسمه بكل فخر، من كلله الله بالهيبية والوقار، جزاك الله عني كل خير يا أبي الغالي

إلى من لا أستطيع وصفها بالكلمات، جنّتي على الأرض، التي لا تكفيها كلمات الشكر، دعواتك كانت نوري، ورضاك كان سبيلي، حفظك الله لي يا أمي الغالية.

إلى أختي "لويزة"، التي كانت لي أمًا ثانية بكل ما تحمله الكلمة من معنى، بحنانها، بدفنها، واحتوانها، شكرًا لك لأنك كنت القلب الكبير الذي ألبأ إليه كلما ضاقت بي الأيام.

إلى أختي "أمينة"، يا صاحبة القلب النقي والمواقف الصادقة، كنت دومًا سنْدًا لي في صمتك، وتشجيعك، واستماعك الدافئ لحكاياتي، فشكرًا من القلب

إلى أختي "هاجر"، شكرًا لدعمك المستمر وتشجيعك الذي كان النور الذي أضاء طريقي نحو النجاح، وشكرًا لأنك دائمًا تُنسي همومي بابتسامتك ومحفظتك

إلى أخي الغالي، الذي كان لي سنْدًا حين مالت بي الأيام، وقوة حين خارت عزيمتي، شكرًا لك على حنانك، ووقوفك إلى جانبي دون تردد، دمت لي أختًا وأمانًا

وإلى زوجته الطيبة، التي كانت الحضور الجميل والكلمة الدافئة في وقت الحاجة، جزاك الله خيرًا على لطفك واهتمامك، وعلى وجودك الداعم الذي لا يُنسى

وإلى زهرات حياتي الصغيرة، وسيم ودانيا، وبراعم الأمل التي تنمو في القلب قبل العين، وإلى لين، فراشة البيت وبهجته، حفظكم الله وبارك في أعماركم.

وإلى صديقتي وبنت عمي الغالية فريال، التي كانت نعم العون ونبع الحنان، لك كل الشكر على وقوفك بجانبتي، وعلى كلماتك التي منحنتني القوة، ووجودك الذي كان البلم في لحظات التعب. لن أنسى دعمك ما حييت

أشكر من أعماق قلبي أختي و صديقتي رميصاء، التي كانت صديقةً وفية قبل أن تكون شريكة في هذا العمل. مشاركتها لي في هذه المذكرة لم تكن مجرد تعاون، بل كانت امتدادًا لصداقة صادقة ودعم لا يُقدَّر بثمن. شكرًا لك على الثقة، والروح الطيبة، والمشاركة التي زادت هذا المشروع قيمةً ومعنىً

إلى صديقاتي المقرّبات، رفيقات الدرب والذكريات، إلى الأخوات اللواتي لم تلدهن أمي، إلى البسمات الدافئة والرفيقات التي جمعتني الحياة بهن، إلى من سرنّ معي في طريق النجاح والخير، إلى مريم، حكمة، نسرين، سارة، خديجة، وأمال، أنتنّ البسمات الصادقة والدعوات الخفية التي رافقتني دون انقطاع، دمتنّ لي قلبًا لا يميل.

إلى كل زميلاتي وزملائي في الدراسة، إلى كل من شاركني لحظات النجاح والتعب، إلى كل الأشخاص الذين أحمل لهم المحبة والتقدير، والحمد لله الكريم، لولاه ما بلغنا هذا المنبر المنير.

نور الهدى

## إهداء

إلى من بلّغ الأمانة، ونصح الأمة، إلى نبيّ الرحمة ونور العالمين، محمد صلى الله عليه وسلم، أزمى الصلاة والسلام  
إلى عمي الراحل رحمه الله ، الذي غاب جسده وبقيت روحه حاضرة في قلبي، دعائي لك لا ينقطع، وذكرك لا تزول"  
إلى جدّي الغالي، سندي الأول وذاكرتي الدافئة، وجودك في حياتي نعمة أحمد الله عليها كل يوم، أطل الله في عمرك وحفظك  
لنا.  
إلى والدي العزيز، منبع القوة والأمان، من علمني أن الخطوة الأولى في الحياة تبدأ بثقة، دمت لي سندًا وفخرًا، وأطل الله  
في عمرك  
إلى أمي الغالية، نبع الحنان ودفء الأيام، دعاؤك كان نوري، ورضائك كان سبيلي، دمت لي جنّة على الأرض، حفظك الله  
وأطل في عمرك  
إلى أخي الكبير بلال، أول من علمني معنى الأخوة، وسندي في كل المواقف، دمت لي أخًا عزيزًا وفخرًا لا يُقدّر بثمن"  
إلى أنيس وجواد، إخوتي الصغيرين، زهرات العمر وضحكة البيت، وجودكما يملأ قلبي فرحًا، وأتمنى لكما دربًا مشرقًا "  
وحياة مليئة بالحب والنجاح  
إلى أختي إكرام، رفيقة القلب والروح، كنت دائمًا الحضان الذي ألوذ إليه، والكلمة الطيبة في كل حين، شكرًا لحنانك ودعمك  
الذي لا يُقدّر بثمن  
إلى ابن عمي هيثم، الذي كان دومًا أكثر من أخ، بحضوره الدائم، ووفقاته الصادقة، ووفائه الذي لا يُقدّر، شكرًا لكونك السند "  
الذي أعتزّ به في كل مراحل حياتي  
إلى هدى، أختي قيل أن تكون شريكة، وصديقتي في الحلم والدرب، وجودك في هذا العمل لم يكن مجرد تعاون، بل كان "  
دفعًا، وصدقًا، وسندًا لا يُنسى. شكرًا لك على قلبك النقي، وعلى روحك التي حملتني في لحظات التعب قبل الفرح، دمت لي دائمًا كما  
أنت... الأصيلة الوفية  
وإلى صديقتي العزيزات، الأخوات اللواتي لم تلهن في الحلم والدرب، أنتنّ النعمة التي لا تُحصى، والأمان الذي سكن القلب بصمت، كنتنّ  
لي دفعًا في برد الأيام، وبهجة في لحظات التعب، دمتنّ لي رفيقات عمر لا يبدهن الزمان  
إلى كل زميلاتي وزملائي في الدراسة، إلى كل من شاركني لحظات النجاح والتعب، إلى كل الأشخاص الذين أحمل لهم المحبة  
والتقدير، والحمد لله الكريم، لولاه ما بلغنا هذا المنبر المنير

رميصاء

## Résumé

Le diagnostic automatique de la maladie de Parkinson à partir de la voix, en s'appuyant sur l'apprentissage profond, est une méthode moderne et prometteuse pour la détection précoce de cette maladie neurologique. Cette approche se caractérise par son caractère non invasif et son faible coût. Plus particulièrement, un système de diagnostic de MP par la voix consiste à classifier des signaux vocaux en classes N (Normale) et AN (Anormale : Parkinsonien), en utilisant des algorithmes de Machine Learning (ML). Dans ce projet, on propose d'appliquer des algorithmes d'apprentissage profond sur les paramètres acoustiques pour cette tâche de diagnostic. La conception d'un tel système se base sur une phase d'apprentissage permettant la modélisation des différentes classes et une phase de reconnaissance permettant la classification du signal d'entrée en une classe N ou AN. Les résultats d'une étude comparative entre les performances du système proposé et celles d'un système basé sur des algorithmes d'apprentissage automatique tels KNN et SVM, nous ont montré l'efficacité du classificateur SVM en modes indépendant et dépendant du locuteur. Cependant, l'algorithme d'apprentissage profond montre une faible précision, causée par le nombre limité des données de la base d'apprentissage.

**Mots-clés :** Diagnostic de la MP, les paramètres acoustiques, Apprentissage profond, CNN, KNN, SVM, Analyse de la voix.

## Abstract

Automatic diagnosis of Parkinson's disease from voice, based on deep learning, is a modern and promising method for early detection of this neurological disorder. This approach is characterized by its non-invasive nature and its low cost. Specifically, a voice-based Parkinson's disease diagnostic system involves classifying vocal signals into two classes: N (Normal) and AN (Abnormal: Parkinsonian), using Machine Learning (ML) algorithms. In this project, we propose to apply deep learning algorithms on acoustic features for this diagnostic task. The design of a Parkinson's disease diagnostic system is based on a training phase, which models the different classes, and a testing phase, which classifies the input signal into either class N or AN. The results of a comparative study between the performance of the proposed system and that of a system based on machine learning algorithms such as KNN and SVM, showed us the effectiveness of the SVM classifier in speaker-independent and speaker-dependent modes. However, the deep learning algorithm shows low accuracy, caused by the limited number of data in the training base.

**Keywords:** Parkinson's disease diagnosis, acoustic features, deep learning, CNN, KNN, SVM, voice analysis.

## الملخص

تشخيص مرض باركنسون تلقائيًا من الصوت، بناءً على التعلم العميق، هو طريقة حديثة وواعدة للكشف المبكر عن هذا الاضطراب العصبي. تتميز هذه الطريقة بطابعها غير التدخلية وتكلفتها المنخفضة. تحديدًا، يتضمن نظام تشخيص مرض باركنسون باستخدام الصوت تصنيف الإشارات الصوتية إلى فئتين (طبيعي وغير طبيعي: باركنسوني)، باستخدام خوارزميات التعلم الآلي (ML). في هذا المشروع، نُقترح تطبيق خوارزميات التعلم العميق على المعلمات الصوتية لهذه المهمة التشخيصية. يعتمد تصميم نظام تشخيص مرض باركنسون على مرحلة تدريب، التي تقوم بنمذجة الفئات المختلفة، ومرحلة التعرف، التي تصنف الإشارة المدخلة إلى فئة إما NR أو AN. أظهرت نتائج دراسة مقارنة بين أداء النظام المقترح وأداء نظام قائم على خوارزميات التعلم الآلي مثل KNN و SVM فعالية مصنف SVM في الوضعين المستقل والمعتمد على المتكلم. في حين تُظهر خوارزمية التعلم العميق دقة منخفضة، ويرجع ذلك إلى محدودية البيانات في قاعدة التدريب.

**الكلمات المفتاحية:** تشخيص مرض باركنسون، استخراج المعلمات الصوتية، التعلم العميق، تحليل الصوت، CNN, KNN,

SVM,

---

## Table des matières

Remerciement.....	i
إهداء .....	ii
إهداء .....	iii
Résumé.....	iv
Abstract.....	iv
المخلص.....	v
Table des matières .....	vi
Liste des tableaux.....	x
Liste des figures.....	xi
Liste des acronymes et abréviations :.....	xii
Introduction Générale.....	1
1.1. Introduction .....	2
1.2. Présentation de la maladie Parkinson.....	2
1.2.1. Définition .....	2
1.2.2. Pathophysiologie .....	2
1.2.3. Etiologie.....	3
1.3. Différentes modalités de diagnostic de la maladie Parkinson.....	4
1.3.1. Electroencéphalographie (EEG).....	4
1.3.2. Magnétoencéphalographie (MEG).....	4
1.3.3. Electrocardiogramme (ECG) .....	4
1.3.4. Imagerie par Résonance Magnétique (IRM) .....	4
1.3.5. Diagnostic par l'écriture .....	5

---

1.3.6. Diagnostic par la voix.....	5
1.4. Modifications de la voix dans MP .....	6
1.4.1. Principales caractéristiques de la voix.....	6
1.4.1.1 Paramètres Acoustiques.....	7
1.4.1.2 Paramètres Spectraux .....	7
1.4.2 MFCC (Mel-Frequency Cepstral Coefficients) .....	7
1.5. Fonctionnement d'un système de diagnostic la MP .....	10
1.6. Exemple de système de diagnostic de la MP utilisé la modalité EMG.....	11
1.7. Conclusion .....	12
2.1. Introduction .....	13
2.2. Brève histoire de l'IA .....	13
2.3. Intelligence artificielle .....	14
2.3.1. Définition.....	14
2.3.2. Intelligence artificielle et son impact sur l'aspect médical .....	14
2.4. Apprentissage automatique.....	15
2.4.1 Définition de L'apprentissage automatique (Maching Learning).....	15
2.4.2. Apprentissage supervisé (Supervised Learning).....	16
2.4.3. Algorithmes d'apprentissage de classification .....	16
2.4.3.1. Algorithme de k plus proches voisins (K Nearest Neighbor KNN).....	17
A. Principe de fonctionnement du KNN.....	17
B. Organigramme de KNN .....	18
2.4.3.2. Algorithme de Séparateurs à Vaste Marge (Support Vector Machine SVM) .....	19
A. Principe de fonctionnement de SVM .....	19
B. Organigramme de l'algorithme SVM .....	20
C. Types de SVM.....	20
2.4.4 Règle de vote .....	22

---

2.5. Apprentissage profond.....	23
2.5.1. Domaines d’application de l’apprentissage profond .....	23
2.5.2. Evolution des réseaux de neurones .....	24
2.5.3. Réseau de Neurones Artificiels.....	25
2.5.3.1. Définition du Réseau de Neurones Artificiels .....	25
2.5.3.2. Types du Réseau de Neurones Artificiels .....	26
2.5.3.2.1 Perceptron multicouches (Multi Layer Perceptron MLP).....	26
2.5.3.2.2 Réseaux de neurones récurrentes (Recurrent neural networks RNN).....	26
2.5.3.2.3 Réseaux de neurones à convolution (Convolutional neural network CNN).....	26
2.6. Etat de l’art sur le diagnostic acoustique de la maladie de parkinson.....	28
2.7. Conclusion.....	29
3.1. Introduction .....	31
3.2. Présentation de la base de données .....	32
3.3. Description de la méthodologie de classification des signaux vocaux (Maching Learning) .....	33
3.4. Méthodologie de classification par 1D-CNN des signaux vocaux (Deep Learning) .....	36
3.5. Implémentation et résultats .....	37
3.5.1. Système basé sur les classificateurs KNN, SVM et 1D-CNN en mode dépendant du locuteur .....	38
3.5.1.1. Système basé sur le classificateur KNN en mode dépendant du locuteur .....	38
3.5.1.2. Système basé sur le classificateur SVM en mode dépendant du locuteur .....	41
3.5.1.3. Système basé sur le modèle 1D-CNN en mode dépendant du locuteur.....	43
3.5.2. Système basé sur les classificateurs KNN, SVM et 1D-CNN en mode indépendant du locuteur .....	45
3.5.2.1. Système basé sur les classificateurs KNN en mode indépendant du locuteur .....	46
3.5.2.2. Système basé sur les classificateurs SVM en mode indépendant du locuteur .....	48
3.5.2.3. Système basé sur le modèle 1D-CNN en mode indépendant du locuteur.....	51

---

3.6. Etude comparatives des performances des classificateurs KNN, SVM et la méthode 1D-CNN en mode dépendant et indépendant de locuteur .....	53
3.7. Conclusion .....	54
Conclusion générale.....	55
Bibliography .....	56

---

## Liste des tableaux

<b>Tableau 2.1.</b> Études Précédentes sur les Performances de Divers Classificateurs Basés sur des Paramètres Vocaux pour le Diagnostic de la Maladie de Parkinson.....	29
<b>Tableau 3.1.</b> Performance du classificateur KNN selon le nombre de voisins et le type de distance (mode dépendant du locuteur).....	39
<b>Tableau 3.2.</b> Performance du classificateur SVM selon le type des noyaux (mode dépendant du locuteur).....	41
<b>Tableau 3.3.</b> Les hyperparamètres clés choisis pour entraîner un réseau de neurones convolutifs unidimensionnel (1D-CNN).....	44
<b>Tableau 3.4.</b> Performance du classificateur KNN selon le nombre de voisins et le type de distance (mode indépendant du locuteur).....	47
<b>Tableau 3.5.</b> Performance du classificateur SVM selon le type des noyaux (mode indépendant du locuteur).....	49
<b>Tableau 3.6.</b> Les hyperparamètres clés choisis pour entraîner un réseau de neurones convolutifs unidimensionnel (CNN 1D). ....	52
<b>Tableau 3.7.</b> Les hyperparamètres clés choisis pour entraîner un réseau de neurones convolutifs unidimensionnel (CNN 1D). ....	54

---

## Liste des figures

<i>Figure 1.1. Location of substance noire et la différence entre normal et la maladie parkinso).</i>	3
<i>Figure 1.2. Diagramme de la méthode MFCC.</i>	8
<i>Figure 1.3. Architecture du système de diagnostic de la MP.</i>	10
<i>Figure 1.4. Schéma fonctionnel du système de classification.</i>	11
<i>Figure 2.1. Organigramme KNN.</i>	19
<i>Figure 2.2. Principe de séparateur a vaste marge.</i>	20
<i>Figure 2.3. Organigramme SVM.</i>	20
<i>Figure 2.4. Principe d'un SVM linéaire.</i>	21
<i>Figure 2.5. Problème d'un SVM non linéaire.</i>	22
<i>Figure 2.6. La relation entre l'intelligence artificielle, le ML et le Deep Learning.</i>	24
<i>Figure 2.7. Un neurone artificiel.</i>	25
<i>Figure 2.8. Architecture générale d'un RNA.</i>	25
<i>Figure 2.9. Exemple d'architecture CNN.</i>	27
<i>Figure 3.1. Schéma du système propose, basé sur le classificateur KNN.</i>	34
<i>Figure 3.2. Schéma du système propose, basé sur le classificateur SVM.</i>	35
<i>Figure 3.3. Schéma du système 1D-CNN.</i>	36
<i>Figure 3.4. Performance TCV du classificateur KNN selon la distance.</i>	39
<i>Figure 3.5. Performance TCS du classificateur KNN selon la distance</i>	40
<i>Figure 3.6. Performance TCV selon le type de noyau (SVM).</i>	42
<i>Figure 3.7. Performance TCS selon le type de noyau (SVM).</i>	42
<i>Figure 3.8. Précision d'entraînement au cours des itérations (mode dépendant).</i>	44
<i>Figure 3.9. Perte d'entraînement au cours des itérations (mode dépendant).</i>	45
<i>Figure 3.10. Performance TCV du classificateur KNN selon la distance.</i>	47
<i>Figure 3.11. Performance TCS du classificateur KNN selon la distance.</i>	47
<i>Figure 3.12. Performance TCV de SVM selon le type de noyau.</i>	49
<i>Figure 3.13. Performance TCS de SVM selon le type de noyau</i>	49
<i>Figure 3.14. Précision d'entraînement au cours des itérations (mode indépendant).</i>	52
<i>Figure 3.15. Perte d'entraînement au cours des itérations (mode indépendant).</i>	52

---

## Liste des acronymes et abréviations :

**MP** : Maladie de parkinson.  
**EEG** : Electro-encéphalographie.  
**EMG** : Magnétoencéphalographie.  
**IRM** : Imagerie par résonance magnétique.  
**MFCC**: Mel-Frequency Cepstral Coefficients.  
**LPC**: Linear predictive coding  
**ECG**: Electrocardiogramme.  
**IA**: Intelligence artificiel.  
**ML**: Machine learning.  
**DL**: Deep learning.  
**KNN**: k-nearest neighbors  
**SVM**: Support vector machine.  
**MLP**: Multi-Layer Perceptron.  
**RNN**: Recurrent neural networks  
**LSTM**: Réseaux long-short term memory.  
**CNN** : Convolutional neural network  
**NR** : Normale.  
**AN** : A Normale.  
**PDA** : Personal Digital Assistant  
**ANN** : artificial neural network  
**RNA** : Réseau de Neurones Artificiels  
**HSO**: Hyperplan Séparateur Optimal  
**FFNN**: Feedforward neural network  
**DT**: Decision Trees  
**NB**: Naive Bayes  
**LR**: Logistic Regression  
**GB**: Gradient Boosting  
**RF**: Random Forests

---

**SGD:** Stochastic Gradient Descent

**XGB:** Extreme Gradient Boosting

**NN:** Neural Network

**LGB:** Light Gradient Boosting

**1D:** 1 Demension

**TCV :** taux de classification de vecteur

**TCS :** taux de classification de signal

**UPDRS:** Unified Parkinsons Disease Rating Scale



# **INTRODUCTION GENERALE**

## Introduction Générale

La maladie de Parkinson est une pathologie neurodégénérative chronique qui affecte non seulement la motricité, mais aussi la parole. Les altérations vocales observées chez les patients en font un outil prometteur pour un diagnostic précoce et non invasif. Les techniques d'apprentissage automatique et profond permettent aujourd'hui d'analyser automatiquement ces signaux vocaux en extrayant des caractéristiques acoustiques révélatrices de la maladie.

Dans ce contexte, notre travail vise à concevoir un système automatique de diagnostic de la MP à partir du signal vocal, en s'appuyant sur des techniques modernes de traitement du signal et d'intelligence artificielle. Pour cela, nous extrayons des paramètres acoustiques tels que les coefficients cepstraux en fréquences Mel (MFCC), qui sont reconnus pour leur efficacité dans la modélisation de la parole humaine. Deux systèmes sont proposés pour la tâche de diagnostic de la MP.

Dans le premier système, des algorithmes classiques de machine Learning, en particulier (KNN) et (SVM), sont appliqués sur des séquences de vecteurs de paramètres MFCC pour classifier des signaux vocaux en deux classes : sujets sains (NR) et patients atteints (AN), en utilisant les descripteurs acoustiques MFCC. Dans le deuxième système, l'algorithme des réseaux de neurones convolutifs (CNN) est utilisé dans le cadre de l'apprentissage profond, afin d'exploiter automatiquement les structures temporelles des signaux vocaux à partir des vecteurs de paramètres MFCC. Ces systèmes proposés fonctionnent en modes dépendant et indépendant du locuteur.

L'objectif est de comparer les performances de ces différents systèmes en termes de précision de diagnostic, de robustesse, et d'identifier la solution la plus adaptée pour un usage clinique.

Ce mémoire est structuré en trois chapitres. Le premier chapitre introduit la MP, ses effets sur la parole, et présente les bases de l'analyse vocale pour le diagnostic, ainsi que la méthode d'extraction des paramètres MFCC. Le deuxième chapitre détaille les algorithmes de classification KNN et SVM, ainsi que les fondements des CNN. Le troisième chapitre présente le système proposé, les différentes étapes d'implémentation, les expériences réalisées ainsi que l'analyse comparative des résultats obtenus.

Nous clôturons ce travail par une conclusion générale résumant l'essentiel des travaux réalisés et traçant quelques perspectives.

## **Chapitre 1**

# **Etat de l'art sur les techniques de diagnostic de la maladie de parkinson**

## **1.1. Introduction**

La (MP) est une affection neurodégénérative du système nerveux. Elle se caractérise par la dégénérescence et la mort progressive de cellules nerveuses dans certaines régions du cerveau. Cela entraîne des troubles de la coordination motrice, accompagnés de symptômes typiques tels que le ralentissement des mouvements, la rigidité musculaire et les tremblements [1] [2]

Dans ce chapitre, nous allons aborder succinctement une présentation de MP, son définition, Pathophysiologie, étiologie et les techniques de diagnostic de cette maladie.

## **1.2. Présentation de la maladie Parkinson**

### **1.2.1. Définition**

La MP est le deuxième trouble neurologique le plus commun après la maladie d'Alzheimer. Cette maladie affecte 0.5 à 1 % de la population entre 65-69 ans et cette incidence augmente de 1 à 3% pour les personnes âgées de 80 ans et plus. De plus, 85% des nouveaux cas sont diagnostiqués chez des gens âgés de 65 ans et plus.

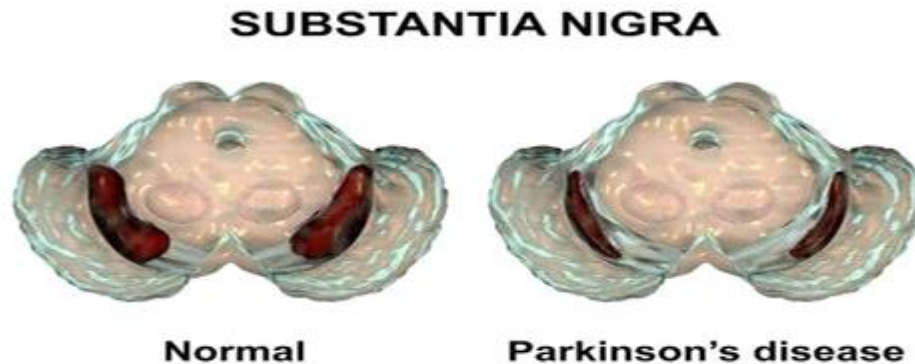
La (MP) est une maladie neurodégénérative dévastatrice caractérisée pathologiquement par la perte de neurones dopaminergiques dans la substance noire, la diminution des taux de dopamine striatale et l'agrégation d'inclusions protéiques intracellulaires, contenant généralement de l' $\alpha$ -synucléine, appelées corps de Lewy. Les signes cliniques classiques de la MP comprennent des tremblements de repos, une rigidité, des troubles de la marche et une bradykinésie. Divers symptômes non moteurs, dont des troubles du sommeil, des troubles de l'humeur, des troubles cognitifs et une démence, sont également fréquemment observés [3].

### **1.2.2. Pathophysiologie**

La MP caractérise par des lésions des neurones (cellules nerveuses) dans la région du cerveau connue sous le nom de substance noire (du latin « matière noire ») présenté dans la (Figure 1.1).

Lorsque les neurones de la substance noire dégèrent, la capacité du cerveau à générer des mouvements corporels est perturbée, produisant des signes et des symptômes caractéristiques de la MP [4].La substance noire est une petite zone située profondément dans le cerveau. On en trouve

une à droite et une à gauche, et l'une est souvent touchée avant l'autre. C'est pourquoi les personnes atteintes de la MP ressentent souvent leurs symptômes principalement d'un seul côté du corps, surtout aux premiers stades [4].



**Figure 1.1.** Location of substance noire et la différence entre normal et la maladie parkinson [5].

### **1.2.3. Etiologie**

On pensait que la MP était principalement causée par des facteurs environnementaux, mais la recherche révèle qu'elle résulte d'une interaction complexe entre génétique et environnement. Ainsi, la MP est aujourd'hui considérée comme une maladie neurodégénérative à progression lente qui débute des années avant que le diagnostic puisse être posé, implique de multiples aires neuroanatomiques, résulte d'une combinaison de facteurs génétiques et environnementaux et se manifeste par une large gamme de symptômes. Cette complexité s'accompagne de défis cliniques. En particulier, il n'existe pas des tests permettant un diagnostic définitif à un stade précoce de la maladie [6].

Concernant les causes environnementales, l'une des conclusions les plus importantes tirées parmi plusieurs études est que l'exposition aux pesticides, en particulier au paraquat et à la roténone, peut être un déclencheur de la MP en induisant une dégénérescence sélective des neurones dopaminergiques dans la substance noire [7].

### **1.3. Différentes modalités de diagnostic de la maladie Parkinson**

Différentes modalités ont été mises en place afin de faciliter le diagnostic de la MP. On peut les résumer comme suit :

#### **1.3.1. Electroencéphalographie (EEG)**

L'EEG dans la (MP) montre un ralentissement généralisé, avec des anomalies distinctes selon la présence de démence. Les patients atteints de démence de Parkinson présentent une augmentation marquée de l'activité delta, tandis que ceux sans démence ont plus d'activité thêta et des fréquences plus lentes [8].

#### **1.3.2. Magnétoencéphalographie (MEG)**

La magnétoencéphalographie (MEG) permet d'analyser l'activité oscillatoire et la connectivité cérébrale dans les maladies neurodégénératives telles que la MP et la maladie d'Alzheimer (MA). Des schémas spécifiques d'altérations ont été identifiés, variant selon la présence ou l'absence de démence. Dans la MP, ces anomalies peuvent réagir différemment au traitement dopaminergique [9].

#### **1.3.3. Electrocardiogramme (ECG)**

La MP entraîne une perturbation de la régulation cardiovasculaire autonome circadienne, comme le démontrent les mesures spectrales de la variabilité de la fréquence cardiaque et la pente de la relation de puissance. Ce défaut semble plus prononcé chez les patients atteints d'une forme plus sévère de la maladie. Le ECG est une méthode utilisée pour détecter ce type de perturbation [10].

#### **1.3.4. Imagerie par Résonance Magnétique (IRM)**

Les chercheurs ont développé des techniques computationnelles en neuroanatomie pour analyser les modifications corticales liées aux symptômes moteurs et non moteurs de la MP. Le (IRM) constitue un outil essentiel pour explorer les altérations anatomiques et fonctionnelles de la connectivité cérébrale chez les patients atteints. Elle permet également de suivre l'évolution de la maladie et d'identifier les changements cérébraux chez les patients précliniques ou à risque, notamment ceux souffrant de troubles de la marche et du sommeil [11]. L'IRM à très faible champ

et la théorie des graphes offrent de nouvelles perspectives pour l'étude des réseaux cérébraux perturbés.

### **1.3.5. Diagnostic par l'écriture**

L'une des caractéristiques cliniques les plus fréquentes de la MP (MP) est la micrographie. Elle se caractérise par une réduction de la taille des lettres et des modifications des aspects cinématiques notamment une augmentation du temps de mouvement, une diminution des vitesses et des accélérations, et une augmentation du nombre de variations de vitesse et d'accélération [12].

L'analyse de l'écriture manuscrites 'est depuis longtemps concentrée sur la reconnaissance automatique du signal d'écriture. En effet, à ce jour, la reconnaissance automatique du signal d'écriture est une technologie performante et mature avec des résultats prometteurs dans plusieurs applications commerciales, notamment dans les services bancaires et postaux pour le traitement des chèques et des adresses], et ce, dans le contexte de l'écriture manuscrite hors ligne, traitée comme une image (acquise sur papier avec un simple stylo). De plus, dans le contexte de l'analyse de l'écriture manuscrite en ligne, cette nouvelle technologie est largement utilisée pour reconnaître les notes prises sur Smartphones, les tablettes ou les assistants numériques personnels (PDA) [13]. Dans cette optique, l'acquisition de l'écriture manuscrite sur tablette graphique numérique est devenue un nouveau domaine de recherche prometteur dans le domaine de la santé, notamment pour la détection précoce des maladies neurodégénératives [13].

### **1.3.6. Diagnostic par la voix**

Les troubles de la parole, comme la dysphonie, sont une manifestation de la dysarthrie parkinsonienne. En effet, la production de la parole implique de nombreux organes et structures cérébrales susceptibles d'être affectés par la MP, entraînant des problèmes et affectant l'articulation et la fluidité. Ces troubles apparaissent plus tardivement, mais altèrent considérablement l'intelligibilité. Décrire le processus de parole permettra d'aborder ses caractéristiques pathologiques [14].

**Remarque :** Bien que les techniques précédemment citées (EEG, ECG, EMG, IRM et diagnostic par l'écriture) soient utilisées dans le diagnostic de la MP, ils nécessitent des

équipements médicaux complexes, un temps d'analyse important, et parfois des conditions spécifiques d'enregistrement. Ceci limite leur utilisation dans un contexte de dépistage rapide ou quotidien, contrairement au diagnostic par la voix. C'est pourquoi nous avons choisi de nous appuyer sur l'analyse de la voix, qui est facile à enregistrer, non invasive, rapide à obtenir, et qui reflète des troubles précoces associés à la MP.

Dans la partie suivante, nous détaillerons les critères de diagnostic basés sur l'analyse de la voix.

## **1.4. Modifications de la voix dans MP**

Les études sur la voix dans MP font souvent référence à la dysarthrie hypokinétique (une réduction de l'amplitude des mouvements des muscles responsables de l'articulation) pour catégoriser les troubles de la voix chez les patients parkinsoniens. Les différentes composantes de la parole affectées par la dysarthrie parkinsonienne sont [15]:

- **Prosodie** : Une perte d'intensité et des modulations de hauteur confèrent à la voix un caractère monotone, le débit est altéré et des troubles de la fluidité sont également observés (palilalie, bégaiement, etc.) ;
- **Articulation** : La précision articulaire des voyelles et des consonnes est altérée.
- **Phonation** : L'intensité de la voix diminue (le patient devient hypo phonique), la hauteur moyenne s'abaisse ou s'élève, la hauteur et l'intensité deviennent instables, et le timbre devient voilé, rauque et voilé ;
- **Rythme** : La capacité à maintenir un rythme de parole constant et instable.

### **1.4.1. Principales caractéristiques de la voix**

L'étude des caractéristiques de la voix permet de mieux comprendre les altérations vocales associées à certaines pathologies. Ces paramètres jouent un rôle clé dans l'évaluation et le diagnostic de troubles tels que la MP. Il existe deux principaux paramètres :

### 1.4.1.1 Paramètres Acoustiques

Les paramètres acoustiques constituent des mesures objectives extraites du signal vocal. Ils permettent d'analyser les altérations de la parole et jouent un rôle essentiel dans l'évaluation de troubles neurologiques comme la MP. Parmi ces paramètres on a :

- **Fréquence fondamentale F0**

Mesurée en Hertz (Hz), elle détermine la hauteur de la voix. Elle dépend du nombre de vibrations par seconde des cordes vocales. Plus la fréquence vocale (F0) est élevée, plus la hauteur de la voix est élevée, et inversement. La mesure de la F0 a surtout un sens dans le suivi du patient, par des mesures répétées, car les normes de F0 dépendent de plusieurs paramètres (sexe, âge, pathologie) [16].

- **Jitter**

Le jitter permet d'objectiver les variations fines de la F0, cycle par cycle (différence entre deux cycles vibratoires contigus) et fait partie des paramètres de stabilité de F0 à court terme.

- **Shimmer**

Le shimmer permet d'objectiver les variations fines d'intensité entre deux cycles vibratoires et fait partie, comme le jitter, des paramètres de stabilité vocale à court terme.

### 1.4.1.2 Paramètres Spectraux

Les paramètres spectraux décrivent la répartition de l'énergie du signal vocal en fonction des fréquences. Ils fournissent des informations précieuses sur la qualité de la voix et permettent de détecter des altérations liées à certaines pathologies, notamment la MP. Parmi ces paramètres on a MFCC (Mel-Frequency Cepstral Coefficients), Méthode de codage prédictif linéaire (LPC) et le Spectrogramme. Dans le cadre de notre projet, nous avons utilisé l'extraction des paramètres MFCC comme base de travail.

## 1.4.2 MFCC (Mel-Frequency Cepstral Coefficients)

La première étape de la reconnaissance vocale consiste à prétraiter le signal vocal et à en extraire les caractéristiques discriminantes. L'extraction de ses caractéristiques poursuit deux

objectifs principaux : représenter le signal vocal sous une forme plus compacte et identifier les caractéristiques vocales capables de discriminer efficacement différentes classes de parole, comme les phonèmes, et d'être insensibles aux facteurs non pertinents pour la reconnaissance vocale [17].

La méthode MFCC proposée par Davis et Mermelstein vise à extraire les paramètres caractéristiques du signal vocal. L'analyse MFCC consiste à exploiter les propriétés du système auditif humain en transformant l'échelle linéaire de fréquences en échelle de Mel, qui fournit la représentation la plus efficace de la voix. Le schéma fonctionnel de (la figure 1.2) suivante décrit succinctement le processus de génération des coefficients MFCC. Maintenant que nous connaissons le fonctionnement général de cette procédure, nous allons expliquer les principaux blocs qui la constituent [18].



Figure 1.2. Diagramme de la méthode MFCC.

### ➤ **Preemphasis**

Preemphasis est une opération de filtrage d'un signal vocal  $\{s_n, n = 1, \dots, N\}$  dans un filtre numérique à réponse impulsionnelle finie du premier ordre dont la fonction de transfert  $H(z)$  est donnée par

$$H(z) = 1 - k z^{-1} \quad (1.1)$$

### ➤ **Segmentation**

Le signal vocal est de nature non stationnaire, alors que les méthodes de traitement du signal utilisent des signaux stationnaires. Il est donc nécessaire, avant d'extraire les paramètres de reconnaissance, de découper le signal en trames de  $N$  échantillons de parole dans l'intervalle. De 10 à 30 ms, cette étape permet d'obtenir, pour chaque segment de parole, un signal quasi stationnaire. Les deux trames adjacentes sont superposées afin d'éviter les transitions brusques entre les trames.

➤ **Windowing**

Les discontinuités aux extrémités des trames sont produites par la segmentation. Le but du fenêtrage est de réduire ces discontinuités du signal en multipliant les échantillons  $\{n=1 \dots, N\}$  de chaque trame du signal vocal par une fenêtre de Hamming. La fenêtre de Hamming est donnée par l'équation suivante.

$$S_n = \left\{ 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right) \right\} S_n \quad (1.2)$$

L'avantage de cette fenêtre est sa résolution en fréquence élevée et son lobe spectral secondaire très petit par rapport à son lobe primaire (atténuation de -43 dB).

➤ **FFT**

Cette phase transforme le signal vocal en un domaine fréquentiel avec la formule, et varie de 0, ... N-1.

$$S_n = \sum_{k=0}^{N-1} S_k e^{-j2\pi \frac{kn}{N}} \quad (1.3)$$

➤ **Mel Filtering with Filter Bank**

La méthode MFCC permet d'extraire des paramètres selon l'échelle Mel dans le domaine fréquentiel. En effet, la perception de la parole par le système auditif humain repose sur une échelle fréquentielle similaire à l'échelle Mel. Cette échelle est linéaire aux basses fréquences et logarithmique aux hautes fréquences et est donnée par l'équation suivante :

$$Mel(f) = 2595 \log_{10}\left(1 + \frac{f}{700}\right) \quad (1.4)$$

➤ **Logarithm / DCT**

Les coefficients MFCC peuvent être calculés directement en appliquant la transformée en cosinus discrète (DCT) des logarithmes des énergies obtenus par les filtres triangulaires M.

$$C_i = \sqrt{\frac{1}{2}} \sum_{j=1}^M m_j \cos\left(\frac{\pi i}{N} (j - 0.5)\right) \quad (1.5)$$

Avec i comme nombre de coefficients à extraire et N comme nombre de filtres triangulaires

➤ **Liftering**

L'ordre supérieur des coefficients cepstraux est trop petit. Pour résoudre ce problème, nous utilisons le lifting afin d'élever le cepstre, ce qui accroît les amplitudes et les rend similaires.

$$C_n = \left(1 + \frac{L}{2} \sin\left(\frac{\pi n}{L}\right)\right) C_n \quad (1.6)$$

## 1.5. Fonctionnement d'un système de diagnostic la MP

Le schéma général de fonctionnement d'un système de diagnostic acoustique de la MP est illustré dans la figure 1.3.

Ce type de système comprend généralement deux phases principales : une phase d'apprentissage et une phase de test. La phase d'apprentissage consiste à modéliser deux classes (NR et AN). Pour cela, chaque signal vocal d'apprentissage est soumis à une étape d'extraction des caractéristiques (par exemple, les coefficients MFCC et leurs dérivées dynamiques), permettant de transformer le signal en une séquence de vecteurs de caractéristiques discriminantes. Ces séquences sont ensuite utilisées pour entraîner les modèles des deux classes.

Lors de la phase de test, les signaux vocaux inconnus subissent le même processus d'extraction de paramètres, puis sont soumis aux classificateurs afin d'être affectés à l'une des deux classes. Les performances du système sont ensuite évaluées à l'aide de métriques telles que le taux de reconnaissance, la précision. Les algorithmes de classification couramment utilisés dans ce contexte sont le KNN, le SVM et les réseaux de neurones convolutifs unidimensionnels (1D-CNN).

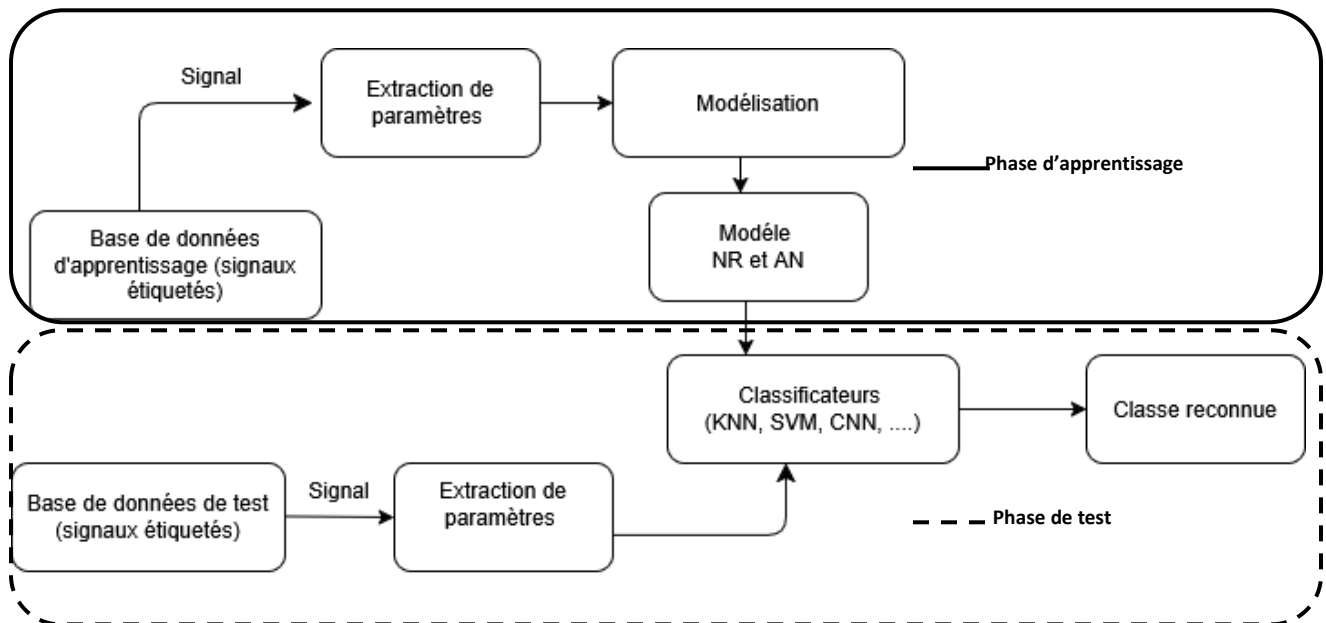


Figure 1.3. Architecture du système de diagnostic de la MP.

## 1.6. Exemple de système de diagnostic de la MP utilisé la modalité EMG

Dans cette étude, les salves d'activité EMG (segments actifs EMG) détectées à l'aide de la technique FM-ALED (Fine Modified Adaptive Linear Energy Detector) (Bengacemi et al., 2020) ont été sélectionnées afin de classer les sujets atteints de MP par rapport aux sujets normaux. Le schéma complet de la tâche est présenté à la figure 1.4.

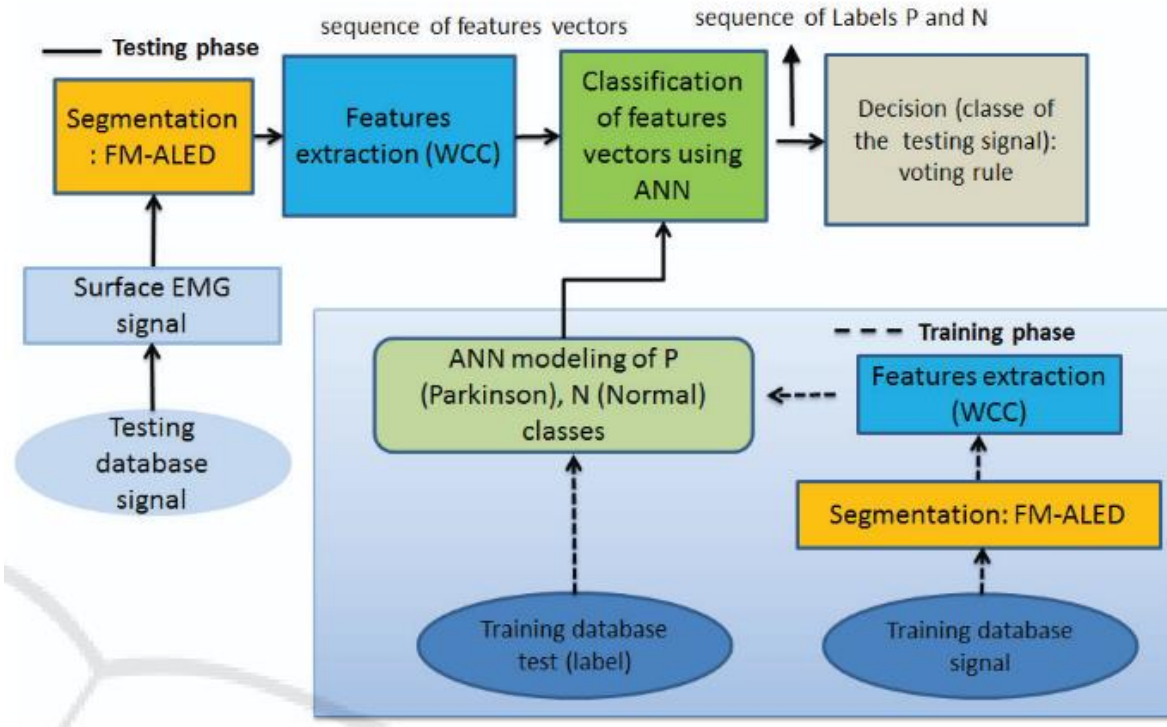


Figure 1.4. Schéma fonctionnel du système de classification [19]

Le système de classification proposé comprend des phases d'apprentissage et de test. La phase d'apprentissage comprend la détection des segments d'activité EMG, l'extraction des caractéristiques et la modélisation des deux classes P et N par la méthode RNA. La phase de test comprend la détection des segments d'activité EMG, l'extraction des caractéristiques, la classification de chaque vecteur de caractéristiques par la technique RNA, puis la classification de la séquence de vecteurs de chaque signal de la base de données de test par la règle de vote afin de déterminer la classe dominante de cette séquence. La base de données des signaux EMG de surface

est issue du projet national français ECOTECH (Buttelli, 2012). Ces signaux EMG ont été obtenus auprès de nombreux sujets (9 sujets sains et 8 sujets atteints de la MP) [19].

## **1.7. Conclusion**

En conclusion, la MP est une pathologie neurologique chronique qui affecte la motricité, l'équilibre et la coordination, en raison d'une diminution progressive de la dopamine. Elle provoque des tremblements, une rigidité musculaire et des troubles de la coordination, impactant la qualité de vie. Ce chapitre a présenté les techniques de diagnostic de la MP, notamment celles basées sur les signaux vocaux. Il ressort que les troubles les plus marqués concernent la prosodie et l'articulation, bien que la phonation et le rythme puissent également être altérés. Les caractéristiques vocales ont été définies comme base pour un système de détection automatique. Le chapitre suivant portera sur un système de diagnostic fondé sur l'apprentissage profond appliqué à l'analyse de la voix ainsi que des classifieurs de machine learning comme KNN et SVM.

## **Chapitre 2**

# **Diagnostic de la maladie de parkinson basé sur l'analyse de la voix combinée avec l'apprentissage profond**

## **2.1. Introduction**

L'intelligence artificielle (IA) regroupe des techniques qui permettent aux machines de simuler certaines capacités humaines comme la perception le raisonnement et la décision. Selon John McCarthy la définit comme « La science de créer des machines intelligentes » [5].

L'apprentissage automatique, né avec Alan Turing, a évolué vers l'apprentissage profond, capable de traiter de grandes quantités de données. Les réseaux de deep learning se distinguent par leur efficacité et leur capacité à extraire automatiquement des caractéristiques complexes [20].

Dans ce chapitre, nous avons présenté un bref historique de l'intelligence artificielle, son évolution et son impact dans le domaine médical. Nous avons également abordé l'apprentissage automatique, ses principaux types, les classifieurs KNN et SVM, ainsi que le deep learning, en retraçant l'histoire des réseaux de neurones et leurs différents types.

## **2.2. Brève histoire de l'IA**

Le terme « intelligence artificielle » a été proposé en 1956 par John McCarthy, dans une loi reprenant les principes de ce domaine naissant. En effet, au cours de l'été de la même année, John McCarthy et Marvin Minsky ont tenu une conférence au Dartmouth College (New Hampshire), réunissant d'autres scientifiques pour discuter des domaines émergents de la cybernétique et de l'informatique. À cette époque, certains chercheurs envisageaient des machines intelligentes uniquement fondées sur des règles logiques, par exemple les systèmes experts (SE). Un exemple est le Logic Theorist, un programme capable de démontrer des théorèmes mathématiques en tant que mathématicien en utilisant des arbres de décision [21].

Dans les années 1950, un courant opposé à l'IA logique a émergé. Il s'appuyait sur les théories de Donal Hebb, psychologue et neurobiologiste canadien, qui étudiait le rôle des connexions neuronales dans l'apprentissage. L'idée principale était de concevoir des machines dont le fonctionnement s'inspirait du cerveau humain ou animal, et qui étaient capables de s'auto-entraîner. Inspirés par les connexions entre les neurones biologiques, ces scientifiques ont modélisé la même architecture dans le Réseau de Neurones Artificiels (ANN), ouvrant une nouvelle ère en ML. Désormais, le neurone artificiel est devenu le protagoniste incontesté au cœur des ANN, car il constitue le fondement du cerveau [21].

En 1988, Yann le Cun proposa le premier réseau neuronal convolutif (CNN) à quatre couches seulement pour la reconnaissance des caractères écrits, toujours inspiré de la structure du cortex visuel. LeNet5, une version ultérieure du premier CNN, fut commercialisée pour lire automatiquement entre 10 et 20 % des chèques déposés aux États-Unis. Malgré ce succès, une nouvelle période sombre commença en 1995 pour les RNA et dura une quinzaine d'années. Les CNN étaient jugés trop complexes et exigeants en termes de calcul. Entre 1992 et 1995, la machine à vecteurs de support (SVM), développée par Isabelle Guyon, Vladimir Vapnik et Bernhard Boser, devint la méthode d'apprentissage automatique privilégiée [21].

## **2.3. Intelligence artificielle**

### **2.3.1. Définition**

Selon le dictionnaire français, l'intelligence artificielle est un ensemble de théories et de techniques mises en œuvre pour créer un système informatique simulant l'intelligence humaine. En d'autres termes, l'IA est un domaine très vaste qui vise à construire des machines intelligentes capables d'effectuer des tâches qui requièrent généralement l'intelligence humaine. Un système est considéré comme intelligent s'il est capable d'apprendre et d'exécuter une tâche liée au processus qui lui a été assigné, sans intervention humaine et avec une grande précision. L'intelligence artificielle a été utilisée dans divers domaines, tels que la finance et la banque, l'armée, la médecine, le renseignement policier, le droit, la logistique, les transports, l'industrie et la robotique [20].

### **2.3.2. Intelligence artificielle et son impact sur l'aspect médical**

L'IA peut avoir un impact positif sur la pratique de la médecine de différentes façons, que ce soit en accélérant le rythme de la recherche ou en aidant les cliniciens à prendre de meilleures décisions. Voici quelques exemples d'utilisation de l'IA dans la Figure 2.1.



**Figure 2.1.** Les domaines d'application de l'IA en médecine [22].

## 2.4. Apprentissage automatique

### 2.4.1 Définition de L'apprentissage automatique (Maching Learning)

L'apprentissage automatique est une branche de L'IA qui permet aux ordinateurs d'apprendre à partir de données sans explicitement programmés (Donner à une machine la capacité d'apprendre sans la programmer de façon explicite).

Contrairement aux méthodes traditionnelles basées sur des équations mathématiques ou des lois physiques, l'apprentissage automatique est utilisé lorsque les problèmes sont complexes et difficiles à formaliser à l'aide de modèles analytiques. Au lieu de programmer explicitement des règles, la machine apprend à partir d'un grand nombre d'exemples et de caractéristiques pertinentes, et les modèles résultants seront capables de prédire les résultats à partir de nouveaux exemples [23].

Il existe trois principaux paradigmes d'apprentissage automatique : Apprentissage par renforcement, apprentissage non supervisé et apprentissage supervisé. Dans la partie suivante, nous

présentons en détail l'approche d'apprentissage supervisé que nous avons utilisée dans le cadre de notre travail.

### **2.4.2. Apprentissage supervisé (Supervised Learning)**

L'apprentissage supervisé est une technique de machine learning qui utilise des jeux de données d'entrée et de sortie étiquetés par des humains pour entraîner des modèles d'intelligence artificielle (IA). Alors, c'est nous qui donnons à la machine des exemples que la machine doit étudier pour en créer ce qu'on appelle un Dataset (c'est un tableau de données). On veut que la machine apprenne la relation qui relie  $x$  et  $y$  selon la fonction  $(f: x \rightarrow y)$  [20].

Dans le cas d'apprentissage supervisé, les classes d'appartenance des données sont connues. La classification supervisée ou la classification inductive a pour objectif de chercher à expliquer et à prédire l'appartenance de documents à des données connues a priori. Ainsi, un classificateur est un dispositif prenant en entrée  $n$  valeurs correspondant aux  $n$  entités  $e_1, e_2, \dots$ , en extraites d'un objet à classifier, et produisant une seule sortie.

Dans le cadre de ce projet de fin d'études, nous nous appuyerons sur l'apprentissage supervisé pour aborder des problématiques de classification. La section suivante sera consacrée à la présentation des principaux algorithmes de classification supervisée issus de la littérature. Il ne s'agit pas de passer en revue toutes les méthodes disponibles, mais de présenter celles, parmi les plus classiques, que nous avons choisies pour ce projet en raison de leurs propriétés adaptées à notre cas d'étude.

### **2.4.3. Algorithmes d'apprentissage de classification**

Les algorithmes d'apprentissage de classification sont des méthodes de l'apprentissage automatique supervisé qui permettent à un système d'apprendre à associer des données d'entrée à des étiquettes de sortie, en s'appuyant sur des exemples déjà connus. Leur objectif principal est de prédire la catégorie ou la classe à laquelle appartient une nouvelle donnée.

Ils sont utilisés dans une grande variété d'applications réelles : détection de spam, diagnostic médical, reconnaissance vocale ou faciale, systèmes de recommandation, ... ,etc [24].

### **2.4.3.1. Algorithme de k plus proches voisins (K Nearest Neighbor KNN)**

Le classificateur **KNN** est couramment utilisé dans le domaine de la reconnaissance de formes, en raison de sa simplicité. Le KNN est un algorithme de classification supervisée qui nécessite une phase d'apprentissage et une phase de test.

Lors de la phase d'apprentissage, chaque classe est représentée par un ensemble de vecteurs de caractéristiques étiquetés, où chaque vecteur correspond à une instance de cette classe.

Dans la phase de test, chaque vecteur de caractéristiques à tester est comparé à l'ensemble des vecteurs étiquetés stockés dans la base d'apprentissage. La classe reconnue est celle qui obtient la majorité des votes parmi les K plus proches voisins [25].

#### **A. Principe de fonctionnement du KNN**

Le principe de fonctionnement de l'algorithme des k plus proches voisins (KNN) est relativement simple. L'algorithme KNN repose sur le principe que des échantillons similaires ont tendance à être regroupés dans des zones similaires de l'espace des caractéristiques. Ainsi, lorsqu'un nouvel exemple est présenté à l'algorithme KNN, celui-ci détermine d'abord les k échantillons les plus proches de l'exemple dans l'espace des caractéristiques.

Ces échantillons sont appelés les "**plus proches voisins**" de l'exemple. La distance entre les échantillons est calculée à l'aide d'une mesure de distance appropriée, telle que la distance euclidienne. Une fois que les k plus proches voisins ont été identifiés, l'algorithme KNN peut alors prédire la **classe** ou la valeur de la variable cible de l'exemple en utilisant une méthode de vote majoritaire. Cela signifie que la classe la plus fréquente parmi les k voisins sera attribuée à l'exemple (voir figure 2.2). Si l'algorithme KNN est utilisé pour la régression, la valeur moyenne des k plus proches voisins sera utilisée pour prédire la valeur de la variable cible. Le choix de la valeur de k est important dans l'algorithme KNN car elle peut avoir un impact significatif sur les performances de l'algorithme. Si k est trop petit, l'algorithme peut être sensible aux bruits de données, tandis que si k est trop grand, l'algorithme peut ne pas être suffisamment flexible pour capturer les variations dans les données. Le choix de la mesure de distance est également important et dépendra des caractéristiques des données utilisées [26]. Il existe plusieurs relations pour calculer les distances, parmi eux :

#### **➤ Distance Euclidienne**

La distance Euclidienne entre les deux vecteurs X et Y est définie par :

$$d_E(X, Y) = \sqrt{\sum_{i=1}^m (x_i - y_i)^2} \quad (2.1)$$

➤ **Distance Cosinus**

Cette distance est définie comme suit :

$$d_{cos}(X, Y) = 1 - \frac{\sum_{i=1}^m x_i y_i}{\sqrt{\sum_{i=1}^m x_i^2} \sqrt{\sum_{i=1}^m y_i^2}} \quad (2.2)$$

➤ **Distance corrélation**

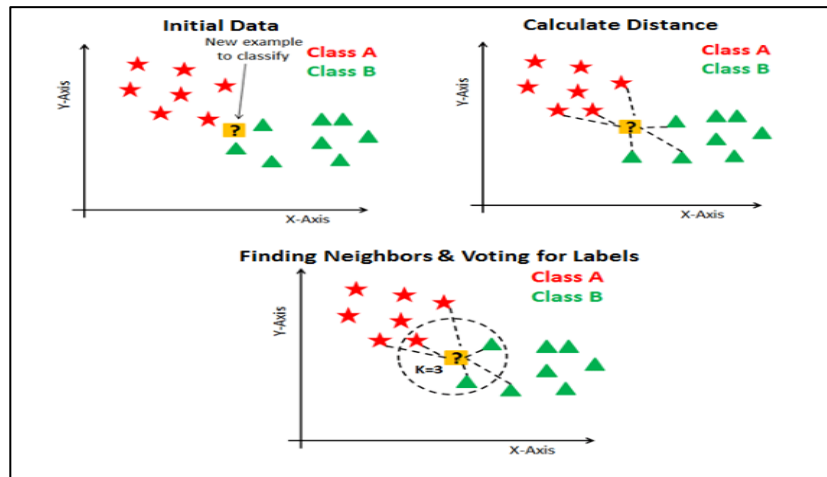
Cette distance est définie comme suit :

$$d_{cos}(X, Y) = 1 - \frac{\sum_{i=1}^m (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^m (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^m (y_i - \bar{y})^2}} \quad (2.3)$$

➤ **Distance Citybock**

Cette distance est définie comme suit :

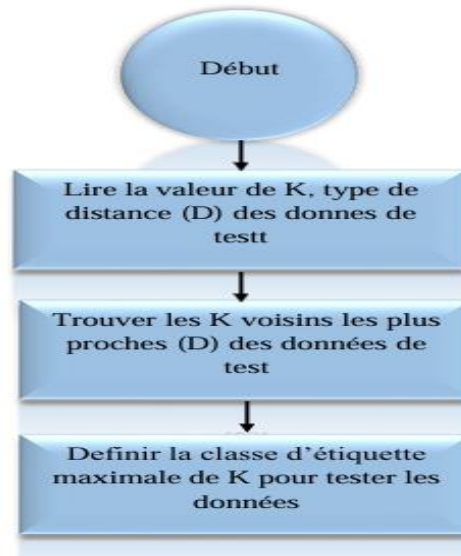
$$d_E(X, Y) = \sum_{i=1}^m |x_i - y_i| \quad (2.4)$$



**Figure 2.2.** Fonctionnement de l'algorithme KNN [26].

**B. Organigramme de KNN**

L'organigramme global de l'algorithme KNN est résumé dans la figure 2.3.



**Figure 2.3.** Organigramme KNN [27]

### **2.4.3.2. Algorithme de Séparateurs à Vaste Marge (Support Vector Machine SVM)**

Les SVM sont des modèles d'apprentissage supervisés. A l'origine, les SVM sont conçus pour résoudre un problème de classification binaire et permettent de construire un classifieur linéaire non probabiliste. Un SVM effectue la classification en construisant un hyperplan N-dimensionnel qui sépare de manière optimale les données en deux catégories [19]

#### **A. Principe de fonctionnement de SVM**

Le but de la méthode SVM est de déterminer si un élément donné appartient à une classe spécifique ou non. Pour cela, on dispose d'un ensemble de données que l'on souhaite diviser en deux groupes distincts.

Le premier groupe comprend les données associées à une certaine classe, ici étiquetées comme "malade". Le second groupe regroupe les éléments qui n'appartiennent pas à cette classe, et qui sont donc étiquetés comme "non malade". L'objectif du SVM est de trouver une frontière optimale permettant de séparer au mieux ces deux ensembles, tout en maximisant la marge entre eux afin d'améliorer la capacité de généralisation du modèle [28]. Le principe de fonctionnement de SVM est illustré dans la figure 2.4.

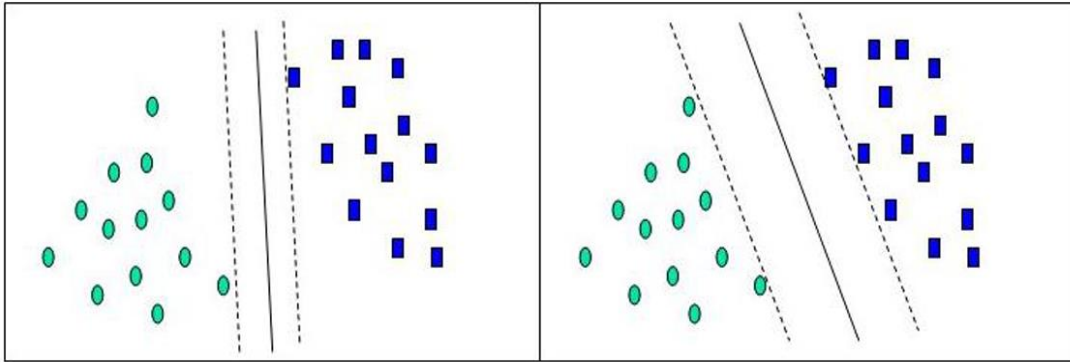


Figure 2.4. Principe de séparateur à vaste marge [19].

### B. Organigramme de l'algorithme SVM

L'organigramme global de l'algorithme SVM est résumé dans la figure 2.5.

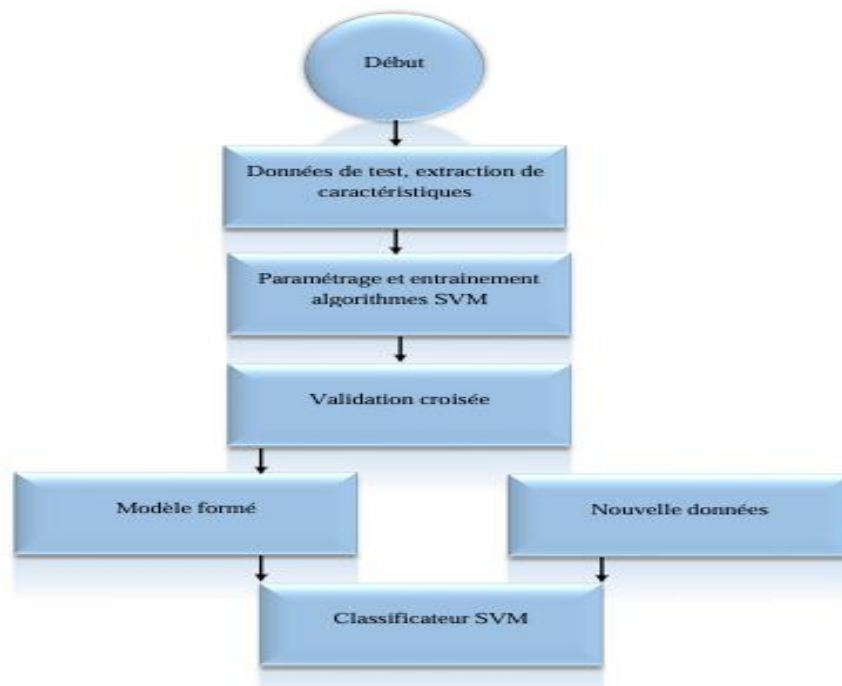


Figure 2.5. Organigramme SVM [27].

### C. Types de SVM

Il existe plusieurs types de SVM selon la nature des données, du problème à résoudre, et de la fonction de séparation souhaitée :

➤ **SVM linéaire**

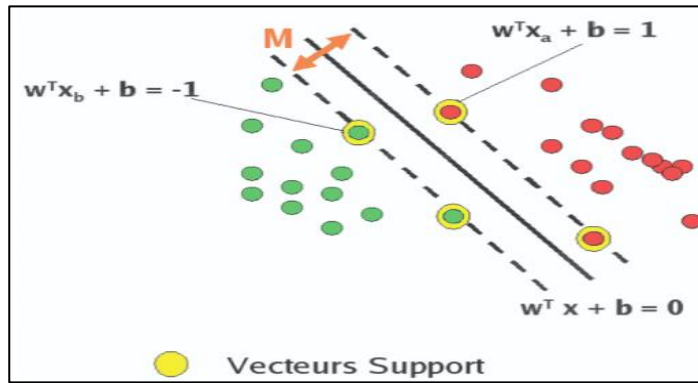
A partir d'un ensemble d'apprentissage  $(\vec{x}_i, \vec{y}_i)$  où  $\vec{x}_i \in \mathbb{R}^m$ , ensemble des données et  $\vec{y}_i \in \{\pm 1\}$ , celui des classes, les SVM consistent à trouver l'Hyperplan Séparateur Optimal (HSO) qui maximise la distance entre l'hyperplan et les deux classes. Cette distance est appelée la marge [29].

L'hyperplan est défini par  $\vec{w} \cdot \vec{x} + b = 0$  ou  $(\vec{w}, b)$  désignent les paramètres de l'hyperplan (respectivement un vecteur normal au plan et le biais). Le classifieur est donné par :

$$f: \vec{x} \in \mathbb{R}^m \rightarrow \text{sign}(\vec{w} \cdot \vec{x} + b) \in \{\pm 1\} \quad (2.5)$$

Néanmoins, ce dernier doit satisfaire :

$$\begin{cases} \vec{w} \cdot \vec{x}_1 + b \geq 0 & \text{si } y_i = +1 \\ \vec{w} \cdot \vec{x}_1 + b \leq 0 & \text{si } y_i = -1 \end{cases}$$



**Figure 2.6.** Principe d'un SVM linéaire [28].

➤ **SVM non linéaire**

Dans la plupart des problèmes, l'hypothèse de linéarité est trop restrictive et le séparateur optimal doit pouvoir prendre une forme plus compliquée. La méthode du noyau est un moyen élégant et efficace pour traiter ce problème : on transfère les données de l'ensemble de départ  $\vec{x} \in \mathbb{R}^m$  vers un ensemble de dimension supérieure  $\phi(\vec{x}_i) \in \mathbb{R}^M$  dans lequel le problème devient séparable linéairement. Le nouveau problème est obtenu en remplaçant les vecteurs de l'ensemble des données qui interviennent seulement dans le produit scalaire  $\vec{x}_i \cdot \vec{x}_j$  par leurs correspondants dans le nouvel espace caractéristique [29].

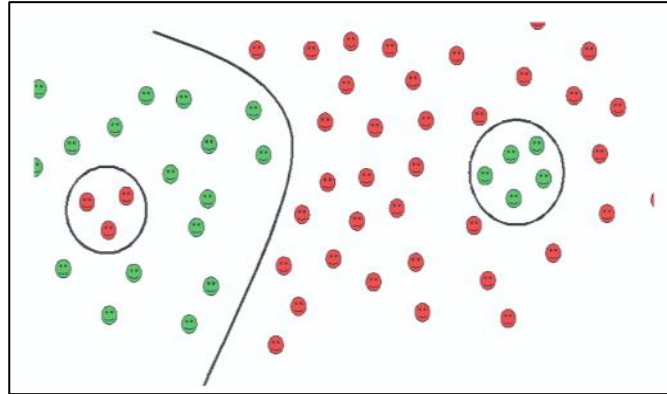
$$\phi(\vec{x}_i) \cdot \phi(\vec{x}_j) \quad (2.6)$$

Ainsi, on introduit la fonction noyau :

$$K(\vec{x}_i, \vec{x}_j) = \phi(\vec{x}_i) \cdot \phi(\vec{x}_j) \quad (2.7)$$

Et la nouvelle fonction de classification est donnée par :

$$f(x) = \sum_{i=1}^N \lambda_i y_i K(\vec{x}_i, \vec{x}) + b \quad (2.8)$$



**Figure 2.7.** Problème d'un SVM non linéaire [28].

#### 2.4.4 Règle de vote

Dans le présent travail, nous proposons un système de diagnostic acoustique de la MP basé sur l'analyse du signal vocal combinée avec l'algorithme de classification KNN, SVM et la stratégie de règle de vote. En fait, ces stratégies de classification avec la règle de vote sont récemment proposées dans plusieurs travaux de recherche tel que [19] [30].

Dans notre travail, les classificateurs **KNN** et **SVM** sont utilisés pour attribuer une classe à chaque vecteur de paramètres extrait du signal vocal, tel qu'un vecteur de caractéristiques MFCC. Chaque signal de test, représenté par une séquence de vecteurs, est ainsi transformé en une séquence d'étiquettes de classes en appliquant l'un de ces deux algorithmes de classification. Pour déterminer la classe finale du signal complet, nous appliquons une stratégie de vote majoritaire sur la séquence d'étiquettes obtenue. Cette méthode consiste à sélectionner la classe la plus fréquemment prédite parmi toutes les trames du signal, ce qui permet de prendre une décision globale robuste concernant la classification du signal vocal.

## **2.5. Apprentissage profond**

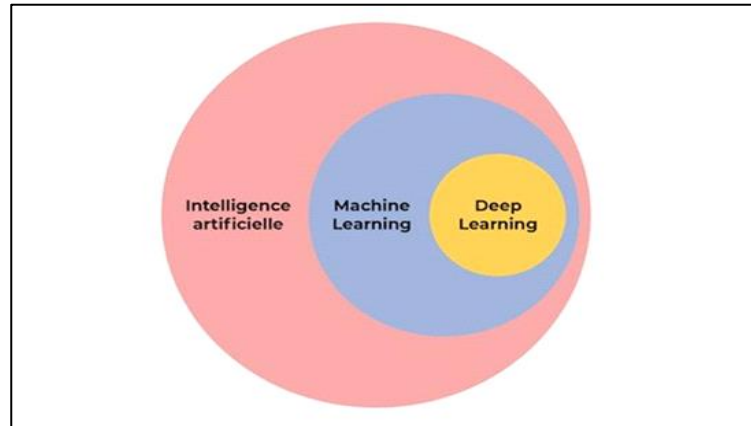
D'après Yann LeCun [4], « L'apprentissage profond ou le Deep Learning permet aux modèles de calcul composé de plusieurs couches d'apprendre des représentations de données avec plusieurs niveaux d'abstraction ». Autrement dit, il est au cœur de l'apprentissage automatique, lui-même une branche de l'IA. Ces techniques apportent des solutions à diverses tâches : classification, régression, segmentation, prédiction, etc [31].

L'apprentissage automatique suppose que les données d'entrée sont structurées, tandis que l'apprentissage profond en extrait automatiquement les caractéristiques discriminantes. Il est utilisé pour traiter des données complexes et variées (texte, audio, images, séries temporelles, etc.). Les méthodes d'apprentissage profond s'appuient sur des couches de neurones artificiels interconnectées pour former un graphe orienté. Ces neurones exécutent une fonction de calcul. Un nombre très élevé de couches justifie le caractère profond de cet apprentissage. Une phase d'apprentissage permet au réseau d'apprendre à effectuer une tâche, qui imite le fonctionnement du cerveau humain. À l'instar des réseaux neuronaux biologiques, les connexions entre les nœuds du graphe symbolisent les synapses [31].

### **2.5.1. Domaines d'application de l'apprentissage profond**

- Reconnaissance des visages.
- Identification selon les mouvements de la marche.
- Lecture des chiffres et des textes manuscrits.
- Reconnaissance et analyses des discours.
- Traduction des langues.
- Contrôler les voitures autonomes.
- Les comportements des robots (position/map estimation).
- Bio-informatique.
- E-commerce (détection de fraude).
- Analyses financières.
- Médecine (diagnostic des maladies).
- Web (les réseaux sociaux) ;

L'apprentissage automatique et le Deep Learning sont des sous-domaines de l'intelligence artificielle qui permettent aux ordinateurs de prendre des décisions intelligentes. En résumé, la figure 2.8 montre la relation entre l'intelligence artificielle, l'apprentissage automatique et le Deep Learning.



**Figure 2.8.** La relation entre l'intelligence artificielle, le ML et le Deep Learning [22].

### 2.5.2. Evolution des réseaux de neurones

Un réseau de neurones artificiels est un système composé d'au moins deux couches : une couche d'entrée et une couche de sortie. Il comprend généralement une ou plusieurs couches intermédiaires appelées *couches cachées* (*hidden layer*). Plus le problème à traiter est complexe, plus le réseau nécessite de couches pour en modéliser les relations. Chaque couche est constituée de nombreux neurones artificiels, chacun spécialisé dans un type de traitement.

Dans ce type de réseau, le neurone artificiel représente l'unité de calcul de base. Il reçoit des entrées issues d'autres neurones ou directement d'une source externe. Cette procédure illustrée dans la Figure 2.9. La sortie du neurone S, appelée activation de sortie, est donc calculée selon l'équation 1 :

$$Y = f(\sum_{i=1}^n wE + b) \quad (2.7)$$

Où  $f$  est la fonction d'activation.

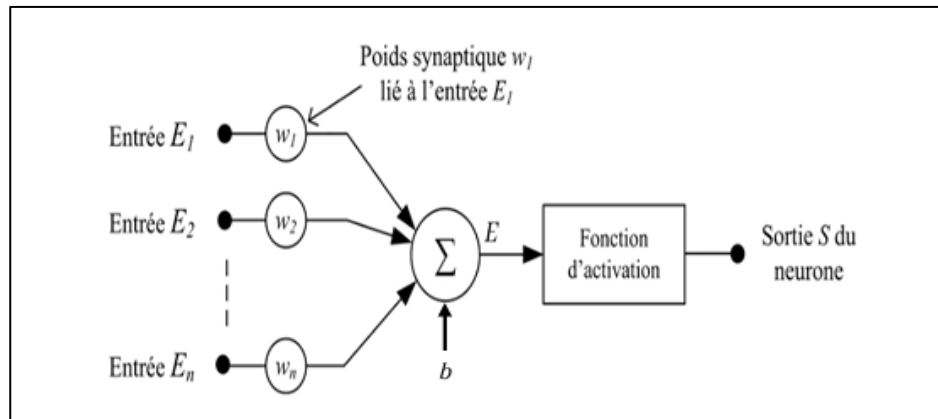


Figure 2.9. Un neurone artificiel [23].

### 2.5.3. Réseau de Neurones Artificiels

Les réseaux de neurones artificiels se déclinent en plusieurs types, chacun étant conçu pour répondre à des tâches spécifiques. Parmi les types qui existe : Réseau de neurone biologique, réseau de neurone formel et le réseau de neurones artificiels.

Dans la section suivante, nous décrivons en détail le réseau de neurones artificiels utilisé dans le cadre de notre travail.

#### 2.5.3.1. Définition du Réseau de Neurones Artificiels

Les réseaux de neurones artificiels se définissent comme des ensembles de processeurs élémentaires fortement interconnectés et opérant en parallèle. Chaque unité de traitement produit une sortie unique à partir des données qu'elle reçoit. Toute organisation hiérarchique de ces réseaux constitue naturellement un réseau [32].

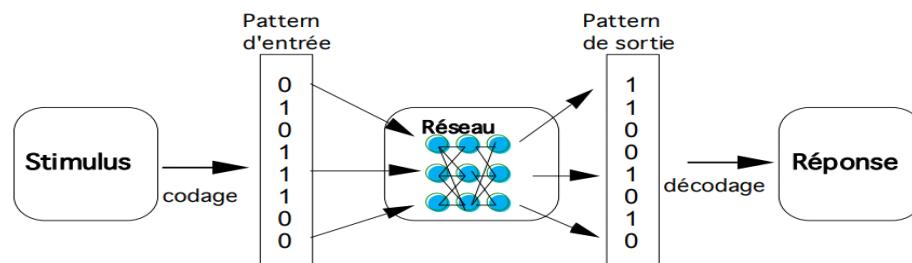


Figure 2.50. Architecture générale d'un RNA [33].

### **2.5.3.2. Types du Réseau de Neurones Artificiels**

Les réseaux de neurones artificiels a plusieurs types qui sont : Perceptron multicouches (Multi Layer Perceptron MLP), Réseaux de neurones récurrentes, Réseaux long-short term memory (LSTM) et le Réseaux de neurones à convolution (CNN). Dans notre travail on a basé sur l'analyse de la voix par l'utilisation de Réseaux de neurones à convolution.

#### **2.5.3.2.1 Perceptron multicouches (Multi Layer Perceptron MLP)**

Les perceptrons multicouches (MLP) sont une architecture de réseau neuronal caractérisée par la présence de couches d'entrée, de couches cachées et d'une couche de sortie. L'information circule de manière unidirectionnelle, de la couche d'entrée vers la couche de sortie. Chaque neurone utilise une fonction d'activation pour introduire des non-linéarités. Les MLP offrent une puissance de calcul élevée grâce à leur architecture multicouche et à la flexibilité apportée par les pondérations et les fonctions d'activation [23].

#### **2.5.3.2.2 Réseaux de neurones récurrentes (Recurrent neural networks RNN)**

Contrairement aux réseaux neuronaux à propagation directe (FFNN), les réseaux neuronaux récurrents (RNN) sont des réseaux neuronaux dont le graphe de connexions contient au moins un cycle. À l'instar des (FFNN), de nombreux types de RNN ont été développés au cours des 30 dernières années, tels que les réseaux d'Elman, de Jordan et les réseaux d'états d'écho [34]. Les réseaux de neurones récurrents (RNN) utilisent généralement des données séquentielles ou des séries temporelles en entrée. Ils sont particulièrement efficaces pour traiter les problèmes impliquant des données ordonnées dans le temps [23].

#### **2.5.3.2.3 Réseaux de neurones à convolution (Convolutional neural network CNN)**

Un CNN peut être considéré comme un perceptron multicouche spécialement conçu pour la reconnaissance de formes, avec un haut degré de stabilité. La translation, la mise à l'échelle, la distorsion et d'autres formes de distorsion sont toutes prises en compte [35].

Les CNN constituent une famille de perceptrons multicouches initialement conçus pour le traitement d'images et la vision par ordinateur. Ils ont notamment montré leur intérêt pour diverses applications, telles que la détection d'objets et la classification d'images [31]. Les CNN sont conçus pour traiter des données sous forme de tableaux de valeurs à N dimensions. Dans ce cas, une image

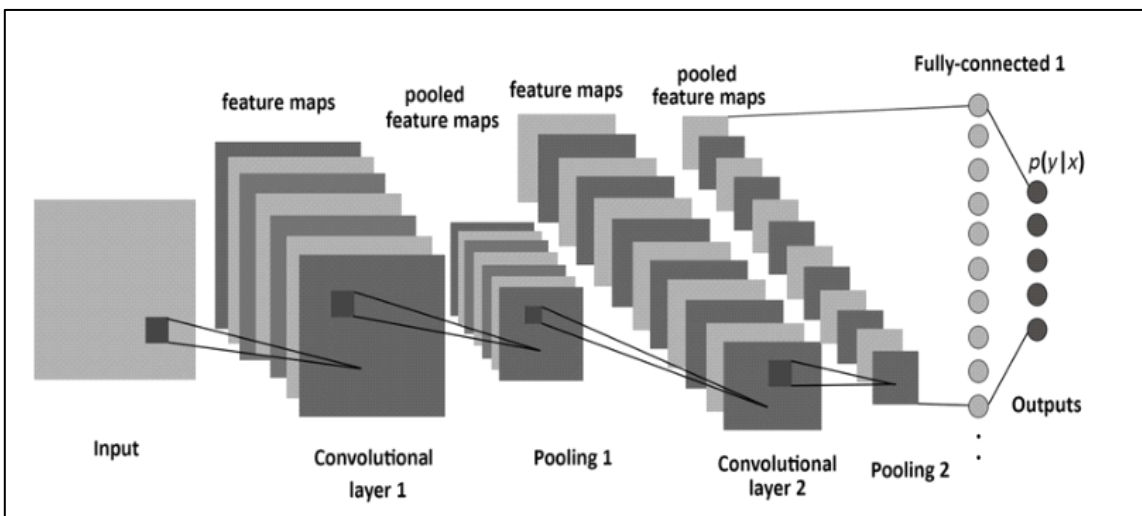
RVB (rouge, vert, bleu) est constituée de trois tableaux 2D, un par canal de couleur. Ces tableaux contiennent les intensités des pixels comme valeurs. Plusieurs autres types de données sont présentés sous forme de tableaux multidimensionnels [31]:

- 1D pour les signaux et les séquences, y compris le langage naturel.
- 2D pour les images ou les spectrogrammes audio.
- 3D pour les vidéos ou les images volumétriques.

Le principe d'un CNN repose sur trois idées clés :

- les connexions locales.
- les pondérations partagées.
- la couche de pooling.

La structure de CNN contient des couches Convolutional, Pooling, ReLU et Fully Connected (Figure 2.11)



**Figure 2.11.** Exemple d'architecture CNN [36].

## 2.6. Etat de l'art sur le diagnostic acoustique de la maladie de parkinson

Ces dernières années, les classificateurs issus de l'apprentissage automatique et de l'apprentissage profond ont été largement exploités dans le cadre du diagnostic de la MP. Le tableau 2.1 présente un résumé de quelques travaux antérieurs.

**Tableau 2.1.** Études Précédentes sur les Performances de divers classificateurs basés sur des Paramètres Vocaux pour le Diagnostic de la Maladie de Parkinson.

Référence	Méthode d'extraction	Classificateur	Base de données	Résultats
[37]	- MFCC -Shimmer et Jitter	-KNN - Decision Trees (DT) -SVM - Naive Bayes (NB) -Logistic Regression (LR) -Gradient Boosting (GB) - Random Forests (RF)	-MDVR-KCL -Italian Parkinson's Voice and Speech database	-Meilleur résultat par classificateur SVM  <b>Accuracy : 98 %</b>
[38]	Analyse de l'intensité et du <b>spectre vocal</b> ; sélection de caractéristiques pour <b>réduire la complexité</b> du jeu de données	-Stochastic Gradient Descent (SGD) -Extreme Gradient Boosting (XGB) -Logistic Regression (LR) - Random Forest (RF) -KNN -Decision Tree (DT)	MDVR-KCL	- <b>SGD</b> : 91 % - <b>XGB</b> : 95 % - <b>Logistic Regression</b> : 91 % - <b>Random Forest</b> : <b>97 %</b> (meilleur) - <b>KNN</b> : 95 % - <b>DT</b> : 95 %
[39]	Caractéristiques vocales (probablement : jitter, shimmer, HNR, MFCC – typiques de la base UCI)	- Support Vector Machine (SVM) - Random Forest (RF) - Decision Tree (DT) - K-Nearest Neighbor (KNN) - Multi-Layer Perceptron (MLP)	UCI Parkinson's Dataset (195 enregistrements vocaux de 31 patients)	<b>MLP</b> : -Accuracy: 98.31 %  <b>SVM</b> : - Accuracy : 95 % - Precision : 98 %
[40]	- Paramètres acoustiques généraux - Composantes temps-fréquence - MFCC - Transformée en ondelettes	- SVM - KNN - Decision Tree (DT) - Neural Network (NN) - Ensemble -Stacking Models	Données vocales de 252 sujets : - 188 patients atteints de PD - 64 sujets sains (contrôles) (base publique, non nommée)	Stacking model (avec Grid Search): -Accuracy: 92.07%

[41]	MFCC	<ul style="list-style-type: none"> <li>- Light Gradient Boosting (LGB)</li> <li>- Extreme Gradient Boosting (XGB)</li> <li>- Random Forest</li> <li>- SVM</li> <li>- KNN</li> <li>- LASSO Regression</li> <li>- Regression logistique</li> </ul>	Parkinson Dataset with Replicated Acoustic Features UCI Machine Learning Repository	-Accuracy:95%
[42]	Features appris automatiquement via 2D-CNN sur Mel-spectrogrammes	CNN profond combinant 2D-CNN + 1D-CNN en architecture end-to-end	<ul style="list-style-type: none"> <li>- Database-1 : Tâches vocales en chinois (voyelle soutenue /a/ et phrase courte).</li> <li>- Database-2 : Tâches vocales en espagnol (voyelles, mots, phrases simples et complexes).</li> </ul>	Database-1 : - Accuracy : 81.6% pour voyelle /a/  Database-2 : - Accuracy :75.3% pour phrase courte.

## 2.7. Conclusion

Dans ce chapitre, nous avons présenté les concepts fondamentaux de l'intelligence artificielle, en retraçant brièvement son historique ainsi que son impact croissant dans le domaine médical. Nous avons ensuite abordé l'apprentissage automatique, en survolant ses principales techniques. Par la suite, nous nous sommes intéressés à l'apprentissage profond, une branche de l'intelligence artificielle en pleine expansion, en mettant en lumière ses domaines d'application les plus pertinents. Enfin, nous avons exposé l'évolution des réseaux de neurones, qui constituent aujourd'hui une avancée majeure dans le développement de solutions intelligentes, et occupent une place centrale dans les systèmes d'intelligence artificielle modernes. Dans le chapitre suivant nous allons à présent aborder la conception d'un système de détection de la MP basé sur l'apprentissage profond, en mettant en œuvre les concepts étudiés précédemment dans un cadre pratique et applicatif.

## **Chapitre 3**

**Conception et implémentation d'un système de diagnostic de la maladie de Parkinson basé sur l'apprentissage profond**

### **3.1. Introduction**

L'objectif de ce chapitre est de présenter en détail la démarche adoptée pour concevoir un système de détection automatique de la MP à partir de signaux vocaux, en s'appuyant sur des techniques d'apprentissage automatique (Machine Learning) et d'apprentissage profond (Deep Learning). Cette phase de conception constitue une étape essentielle pour traduire les objectifs théoriques en une architecture fonctionnelle, structurée et reproductible.

Dans un premier temps, nous décrivons l'architecture générale du système proposé, en détaillant les différentes étapes de traitement du signal, depuis l'acquisition et le prétraitement des enregistrements vocaux jusqu'à l'extraction des caractéristiques pertinentes. Une attention particulière sera portée aux techniques d'extraction comme les coefficients cepstraux en fréquences Mel (MFCC), qui permettent de représenter efficacement le contenu spectral de la voix.

Nous présenterons ensuite les modèles de classification utilisés, en distinguant les approches traditionnelles (telles que SVM, KNN) des modèles de Machine Learning, notamment les réseaux de neurones convolutifs (CNN), qui sont particulièrement adaptés à l'analyse de données spectrales ou de représentations d'images comme les spectrogrammes. L'architecture des modèles CNN, les fonctions d'activation, les stratégies de régularisation, ainsi que les méthodes d'optimisation des hyperparamètres y seront discutées.

- Comment peut-on concevoir un système fiable de détection de la MP à partir de la voix ?
- Pourquoi les coefficients MFCC sont-ils particulièrement adaptés à la représentation des signaux vocaux dans ce contexte ?
- Quelles sont les différences méthodologiques entre les approches traditionnelles de Machine Learning (KNN, SVM) et les approches profondes comme les CNN ?
- Quels sont les critères à considérer pour choisir les hyperparamètres du modèle CNN (nombre de couches, taille du noyau, taux d'apprentissage, etc.) ?
- Quels sont les avantages et les limites de chaque classificateur utilisé dans ce projet (KNN, SVM, CNN) ?

Ainsi, plusieurs expériences sont menées pour répondre sur ces différentes questions. L'implémentation du système de diagnostic et les différents classificateurs est effectuée sous l'environnement Matlab.

### 3.2. Présentation de la base de données

Dans cette étude, nous utilisons la base de données Italian Parkinson's Voice and Speech - 1 [43], développée dans le but d'explorer les altérations de la parole associées à la MP. Cette base comprend des enregistrements de phonation soutenue, de lecture de textes, ainsi que de répétition de phrases, réalisés à la fois par des patients atteints de la MP et par des sujets témoins sains.

Chaque fichier audio est accompagné de métadonnées cliniques telles que l'âge, le sexe, et le score UPDRS (Unified Parkinson's Disease Rating Scale), permettant d'étudier les corrélations entre l'état clinique du patient et les caractéristiques acoustiques extraites de sa voix.

Les fichiers sont enregistrés au format WAV (PCM 16 bits, 44,1 kHz), garantissant une qualité suffisante pour des analyses acoustiques de haute précision.

Deux modes d'évaluation ont été envisagés dans notre étude :

**Mode dépendant du locuteur** : les locuteurs présents dans la phase d'apprentissage sont également présents dans la phase de test. Ce mode permet d'évaluer les performances du système dans un cadre contrôlé.

**Mode indépendant du locuteur** : les locuteurs de l'ensemble de test sont complètement distincts de ceux de l'ensemble d'apprentissage. Ce mode est plus rigoureux et reflète mieux les performances du système dans des situations réelles, où les patients à diagnostiquer n'ont jamais été rencontrés durant l'apprentissage.

Dans le cadre de notre étude, cette base de données a été divisée en deux ensembles :

**Phase d'apprentissage** (mode dépendant) : contient 250 enregistrements vocaux, soit 110 enregistrements de sujets sains (classe normale) et 140 enregistrements de patients parkinsoniens (classe anormale).

**Phase de test** (mode dépendant) : contient 245 enregistrements vocaux, avec également 110 enregistrements de sujets sains et 135 enregistrements de patients atteints de la MP.

Ainsi, la première classe (250 enregistrements, soit environ 50.5 % de la base totale) est utilisée comme base d'apprentissage, tandis que la seconde classe (245 enregistrements, soit environ 49.5 %) est utilisée comme base de test.

**Phase d'apprentissage** (mode indépendant) : contient 260 enregistrements vocaux, soit 110 enregistrements de sujets sains (classe normale) et 150 enregistrements de patients parkinsoniens (classe anormale).

**Phase de test** (mode indépendant) : contient 235 enregistrements vocaux, avec également 110 enregistrements de sujets sains et 125 enregistrements de patients atteints de la MP.

Ainsi, la première classe (260 enregistrements, soit environ 52.53 % de la base totale) est utilisée comme base d'apprentissage, tandis que la seconde classe (235 enregistrements, soit environ 47.47 %) est utilisée comme base de test.

### **3.3. Description de la méthodologie de classification des signaux vocaux (Maching Learning)**

Dans notre étude, nous avons appliqué des algorithmes de classification supervisée, notamment le K-Nearest Neighbors (KNN) et le Support Vector Machine (SVM), afin de distinguer entre les patients atteints de la MP et les sujets sains, à partir de l'analyse de leur voix. Le processus se déroule en deux phases principales : la phase d'apprentissage (training) et la phase de test (testing). Dans un premier temps, les signaux vocaux sont prétraités, puis convertis en une représentation numérique à l'aide de l'extraction des coefficients cepstraux en fréquences de Mel (MFCC).

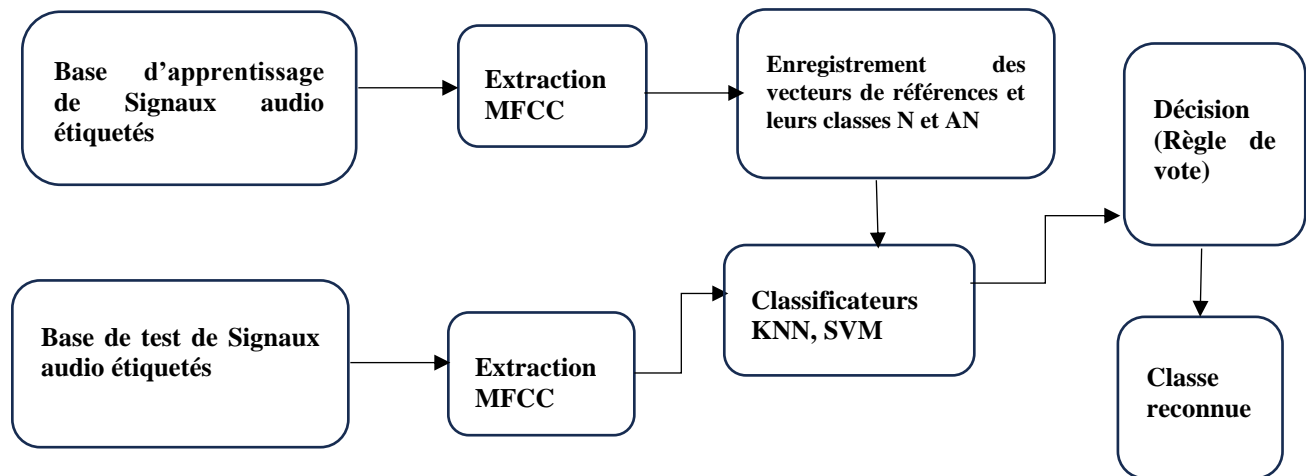
Pour chacune de ces phases, le signal vocal est transformé en une séquence de vecteurs de caractéristiques MFCC, extraits à l'aide d'un outil HTK (Hidden Markov Model Toolkit), couramment utilisé pour l'analyse de la parole. Ces coefficients servent ensuite de vecteurs de caractéristiques pour chaque trame du signal.

Durant la phase d'apprentissage, les modèles KNN et SVM sont entraînés à partir des vecteurs MFCC extraits des signaux d'une base de données étiquetée. Ensuite, lors de la phase de test, les mêmes caractéristiques sont extraites de nouveaux signaux, et les modèles précédemment appris sont utilisés pour prédire la classe (patient ou sain) de chaque vecteur.

Afin d'obtenir une décision globale pour chaque signal vocal complet, une règle de vote majoritaire est appliquée sur les prédictions des différentes trames. Ainsi, un enregistrement est classé comme appartenant à un patient atteint de Parkinson si la majorité des trames sont classées dans cette catégorie, et inversement pour les sujets sains.

Cette approche a permis de comparer les performances des deux classificateurs (KNN et SVM) sur des signaux vocaux réels, en mettant en évidence les avantages de l'extraction MFCC

dans le cadre du diagnostic automatique de la MP. Figure 3.1 montre le Schéma du système propose, basé sur les classificateurs KNN, SVM.



**Figure 3.1.** Schéma du système propose, basé sur les classificateurs KNN, SVM.

Les différentes étapes de fonctionnement de système proposé, basés sur l'utilisation des classificateurs KNN et SVM associée à la stratégie de la règle de vote, sont les suivantes :

➤ **Préparation des données**

La base de données Italian Parkinson's Voice and Speech Database contient des enregistrements vocaux de patients atteints de la MP et de sujets sains. L'objectif de cette préparation est de constituer deux sous phases tels que :

**Phase d'apprentissage** : pour apprendre aux algorithmes les caractéristiques des voix saines et malades.

**Phase de test** : pour évaluer les performances du système.

La répartition des signaux se fait de manière équilibrée pour garantir la représentativité des deux classes.

➤ **Extraction des paramètres acoustiques**

Chaque enregistrement vocal est découpé en trames (~25 ms). Pour chaque trame, on extrait des coefficients MFCC à l'aide de l'outil HTK (Hidden Markov Model Toolkit), qui simule la perception humaine du son. Les MFCC capturent les caractéristiques spectrales de la voix qui sont souvent altérées chez les patients atteints de Parkinson (dysphonie, tremblements, rigidité vocale...). Chaque signal vocal est transformé en une matrice de vecteurs MFCC, représentant le contenu spectral du signal dans le temps.

➤ **Classification des vecteurs MFCC**

Chaque vecteur MFCC extrait d'une trame est soumis à un classificateur supervisé afin de déterminer s'il provient d'une voix saine ou d'une voix parkinsonienne. Dans cette étude, deux classificateurs sont utilisés :

→ **KNN**

Cet algorithme classe chaque vecteur en fonction des classes des K vecteurs d'apprentissage les plus proches dans l'espace des caractéristiques.

Il s'agit d'une méthode simple et efficace, basée sur la distance, qui ne nécessite pas d'apprentissage préalable complexe.

→ **SVM**

SVM construit un hyperplan optimal séparant les données des deux classes dans un espace à dimensions éventuellement élevées (via des noyaux).

Il est particulièrement adapté aux problèmes de classification binaire avec un bon compromis entre complexité et performance, notamment pour des données non linéairement séparables.

➤ **Décision globale par vote majoritaire**

Les prédictions effectuées au niveau des trames peuvent être bruitées et peu fiables individuellement. Pour améliorer la robustesse du diagnostic :

Pour chaque enregistrement vocal, on récupère la classe prédite de chaque trame par le classificateur (KNN ou SVM).

On applique une règle de vote majoritaire : la classe la plus fréquemment prédite sur l'ensemble des trames est attribuée à l'enregistrement complet.

### 3.4. Méthodologie de classification par 1D-CNN des signaux vocaux (Deep Learning)

Dans cette approche, nous utilisons les réseaux de neurones convolutifs unidimensionnels (1D-CNN) pour classifier les enregistrements vocaux en voix saine ou voix parkinsonienne, en se basant sur les vecteurs MFCC extraits du signal audio. La figure 3.2 montre le Schéma du système 1D-CNN.

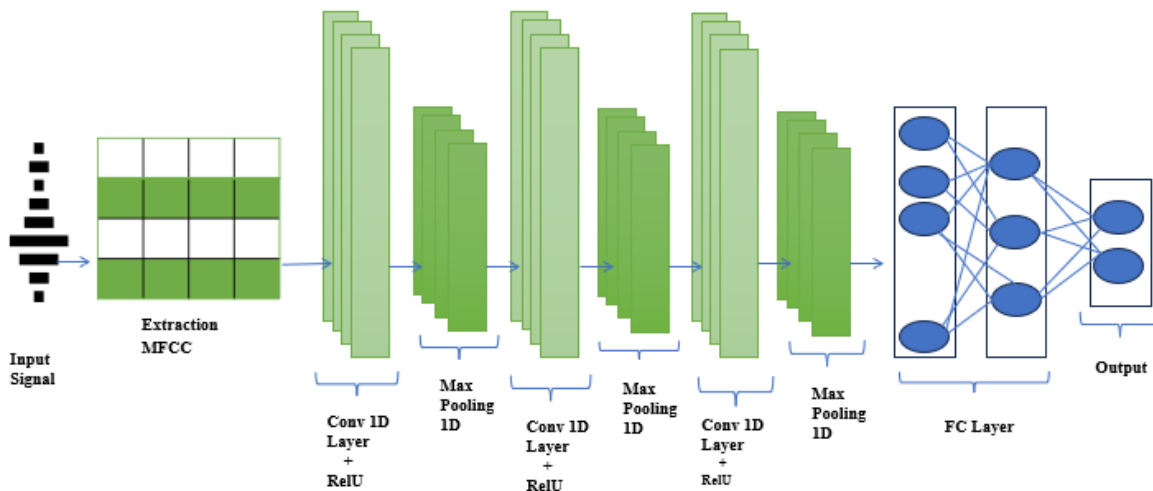


Figure 3.2. Schéma du système 1D-CNN.

Le système d'analyse fondé sur le deep learning, conçu pour le diagnostic de la MP, se déroule en plusieurs étapes clés, détaillées ci-dessous :

#### ➤ **Prétraitement des signaux**

Chaque enregistrement vocal est d'abord segmenté en trames, puis des coefficients MFCC sont extraits pour chaque trame, formant une séquence temporelle de vecteurs MFCC.

Cette représentation 1D (série temporelle de vecteurs) est bien adaptée à l'analyse par 1D-CNN, car elle permet de capturer les motifs temporels caractéristiques des anomalies vocales liées à la MP.

#### ➤ **Construction du modèle 1D-CNN**

Le modèle 1D-CNN est composé de couches convolutives unidimensionnelles qui opèrent sur la dimension temporelle des vecteurs MFCC, suivies de :

- Fonctions d'activation ReLU,
- Couches de pooling pour la réduction de dimension,
- Puis de couches entièrement connectées menant à la classification finale.

Ce réseau apprend automatiquement à extraire des caractéristiques discriminantes entre les voix saines et parkinsoniennes à partir de la dynamique temporelle des coefficients MFCC.

#### ➤ **Entraînement du modèle**

Le 1D-CNN est entraîné à partir d'un ensemble de séquences MFCC labellisées (voix saine ou parkinsonienne).

Une fonction de perte telle que l'entropie croisée est utilisée, et l'optimisation se fait à l'aide d'un algorithme comme Adam.

Les paramètres du réseau sont ajustés afin de maximiser les performances de classification sur les données d'apprentissage.

#### ➤ **Prédiction et décision**

Lors de la phase de test, chaque enregistrement vocal est traité pour extraire ses MFCC, qui sont ensuite transmis au 1D-CNN.

- Le modèle prédit la probabilité d'appartenance à chaque classe (sain/parkinsonien).
- La classe avec la probabilité maximale est attribuée à l'enregistrement.

### **3.5. Implémentation et résultats**

Dans cette section, nous présentons les différentes expériences menées afin de répondre aux diverses questions posées en début de chapitre. Les premières expérimentations sont basées sur l'utilisation des classifieurs **KNN** et **SVM**, appliqués dans deux configurations : en mode dépendant du locuteur et en mode indépendant du locuteur.

Par la suite, une approche par apprentissage profond a été explorée à l'aide d'un réseau de neurones convolutif unidimensionnel (1D-CNN). Ce modèle a été conçu pour exploiter automatiquement les structures temporelles et spectrales des paramètres acoustiques (MFCC), sans avoir recours à une phase explicite de sélection de caractéristiques.

### 3.5.1. Système basé sur les classificateurs KNN, SVM et 1D-CNN en mode dépendant du locuteur

Dans cette section, nous cherchons à identifier la meilleure configuration des classificateurs KNN, SVM et du modèle de réseau de neurones convolutif unidimensionnel (CNN-1D) afin d'obtenir les meilleurs taux de classification des vecteurs (TCV) et le taux de classification des signaux (TCS) pour KNN et SVM et d'obtenir le meilleur taux de reconnaissance pour CNN-1D en mode dépendant du locuteur.

#### 3.5.1.1. Système basé sur le classificateur KNN en mode dépendant du locuteur

L'objectif est de déterminer la configuration optimale du classificateur KNN, en choisissant le nombre de voisins K et le type de distance les plus appropriés, afin d'améliorer la performance du système de diagnostic de la MP en mode dépendant du locuteur, en maximisant les taux de classification (TCV) et (TCS). Alors que les différents types de distances utilisées sont : Euclidean, City block, Cosine et Correlation.

Le tableau ci-dessus (3.1) présente les résultats de performance d'un classifieur KNN appliqué au diagnostic de la MP à partir de la voix, en testant différentes valeurs de **K** (de 1 à 25) et différentes mesures de distance. Deux métriques sont rapportées :

**TCS (%)** : Taux de classification des signaux.

**TCV (%)** : Taux de classification des vecteurs.

**Tableau 3.1.** Performance du classificateur KNN selon le nombre de voisins et le type de distance (mode dépendant du locuteur).

Distance K	Euclidean		Cityblock		Cosine		Correlation	
	TCV	TCS	TCV	TCS	TCV	TCS	TCV	TCS
1	97.19	99.54	96.93	99.54	96.92	99.54	96.94	99.54
2	97.22	99.54	96.94	99.54	97.06	100	97.10	100
3	96.86	99.54	96.66	99.54	96.53	99.54	96.53	99.54
4	96.87	99.54	96.65	99.54	96.60	99.54	96.62	99.54
5	96.75	99.54	96.59	99.54	96.39	99.54	96.38	99.54
6	96.75	99.54	96.58	99.54	96.43	99.54	96.43	99.54
7	96.71	99.54	96.56	99.54	96.32	99.54	96.28	99.54
8	96.71	99.54	96.54	99.54	96.36	99.54	96.33	99.54
9	96.68	99.54	96.54	99.54	96.26	99.54	96.20	99.54
10	96.68	99.54	96.52	99.54	96.29	99.54	96.24	99.54
11	96.66	99.54	96.51	99.54	96.23	99.54	96.17	99.54

12	96.65	99.54	96.50	99.54	96.25	99.54	96.19	99.54
13	96.66	99.54	96.50	99.54	96.21	99.54	96.13	99.54
14	96.65	99.54	96.49	99.54	96.22	99.54	96.15	99.54
15	96.65	99.54	96.49	99.54	96.16	99.54	96.10	99.54
16	96.64	99.54	96.48	99.54	96.18	99.54	96.12	99.54
17	96.63	99.54	96.47	99.54	96.18	99.54	96.08	99.54
18	96.63	99.54	96.46	99.54	96.16	99.54	96.09	99.54
19	96.62	99.54	96.46	99.54	96.14	99.54	96.06	99.54
20	96.63	99.54	96.47	99.54	96.18	99.54	96.11	99.54
21	96.63	99.54	96.46	99.54	96.14	99.54	96.08	99.54
22	96.62	99.54	96.46	99.54	96.15	99.54	96.08	99.54
23	96.61	99.54	96.46	99.54	96.13	99.54	96.05	99.54
24	96.61	99.54	96.45	99.54	96.13	99.54	96.06	99.54
25	96.60	99.54	96.45	99.54	96.11	99.54	96.03	99.54

Ces résultats sont illustrés dans les figures 3.3 et 3.4.

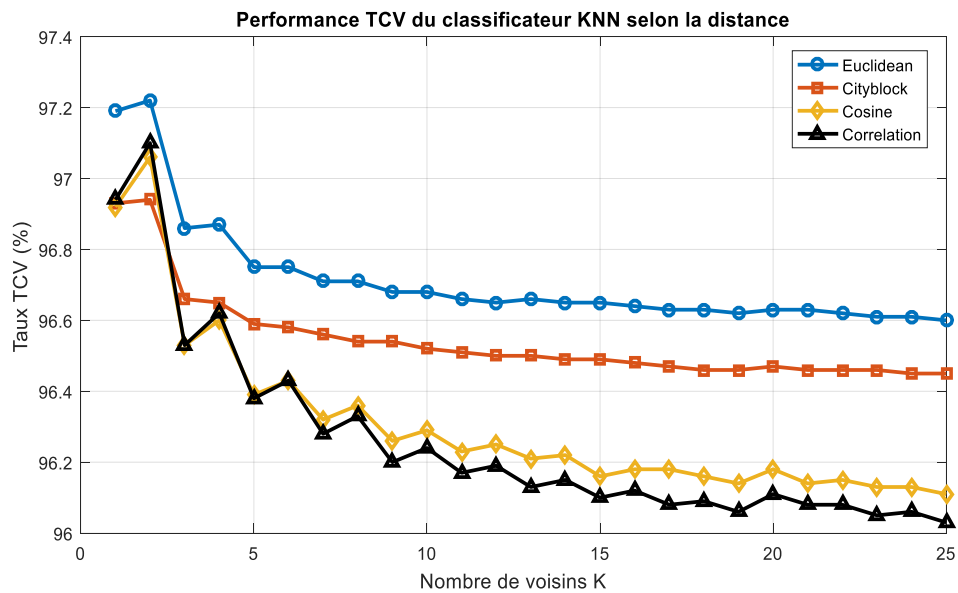


Figure 3.3. Performance TCV du classificateur KNN selon la distance.

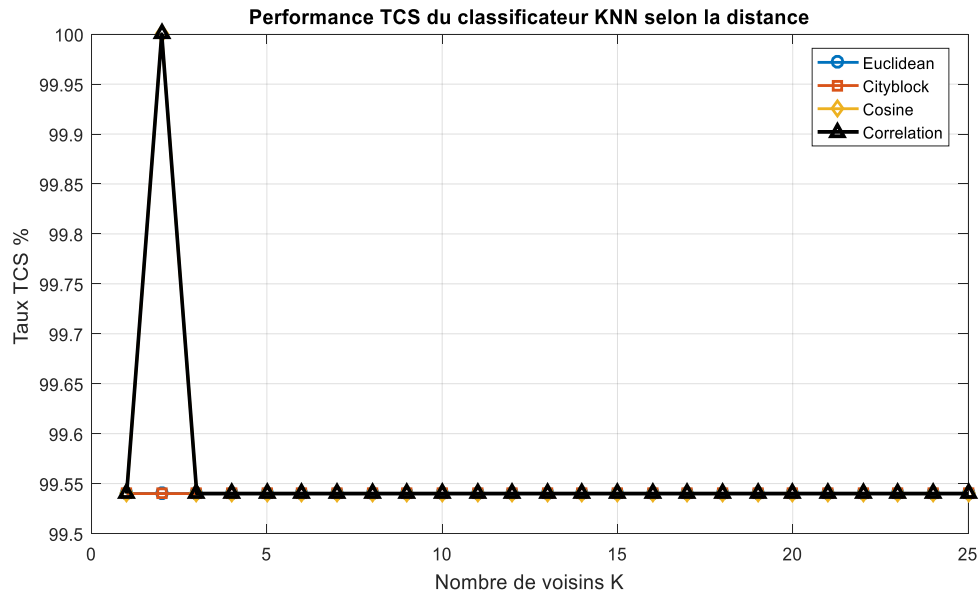


Figure 3.4. Performance TCS du classificateur KNN selon la distance

A partir des résultats illustrés sur le tableau 3.2 ainsi que les figures 3.19 et 3.20, nous pouvons remarquer les points suivants :

- Le système affiche un Taux de classification des signaux (TCS) très élevé et stable (~99.54%), prouvant l'efficacité du vote majoritaire pour compenser les variations locales et assurer une décision globale fiable.
- Les meilleurs Taux de classification de vecteurs (TCV) sont obtenus avec un petit nombre de voisins (K=1 ou K=2), notamment 97.19% avec la distance Euclidienne et même 100% de TCS avec la Corrélation à K=2, indiquant des données bien séparées.
- Le TCS soit stable pour toutes les distances, le TCV varie, les distances comme la Corrélation et la Cosinus sont bien adaptées aux données vocales, tandis que l'Euclidienne reste très performante.
- Le TCV diminue légèrement avec l'augmentation de K mais reste au-dessus de 96%, démontrant une bonne stabilité du classifieur, même si un K trop élevé peut légèrement réduire la précision locale.

### 3.5.1.2. Système basé sur le classificateur SVM en mode dépendant du locuteur

Dans cette partie, nous avons utilisé le classificateur SVM, associé à la règle de vote, afin de déterminer les paramètres TCV et TCS pour le diagnostic acoustique de la MP.

Le tableau 3.2 et les figures 3.5-3.6 résument les résultats obtenus pour un système de classification basé sur le classificateur SVM en mode dépendant du locuteur, c'est-à-dire que le modèle est entraîné et testé individuellement pour chaque locuteur.

**Tableau 3.2.** Performance du classificateur SVM selon le type des noyaux (mode dépendant du locuteur).

Order	Accuracy (%)					
	Polynomial		Linear		Rbf	
	TCV	TCS	TCV	TCS	TCV	TCS
1	92.85	89.40	92.82	91.24	96.12	100
2	94.13	99.07			96.14	100
3	95.36	100			96.16	100
4	95.83	100			96.14	100
5	95.99	100			96.15	100
6	81.07	62.21			96.12	100
7	81.06	62.21			96.16	100

Les figures 3.5 et 3.6 présentent l'évolution des performances du classificateur SVM selon différents paramètres.

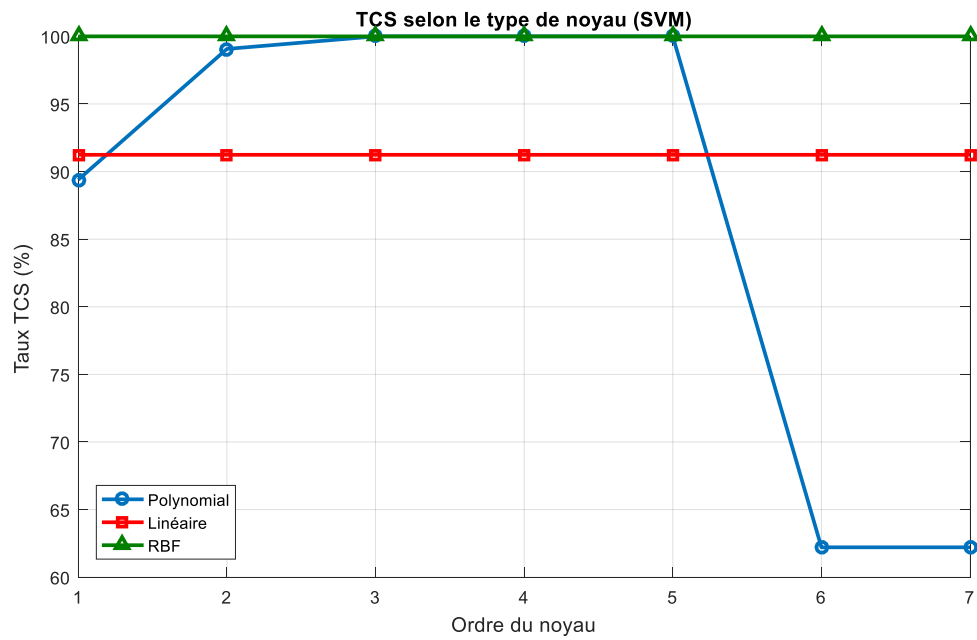


Figure 3.5. Performance TCV selon le type de noyau (SVM).

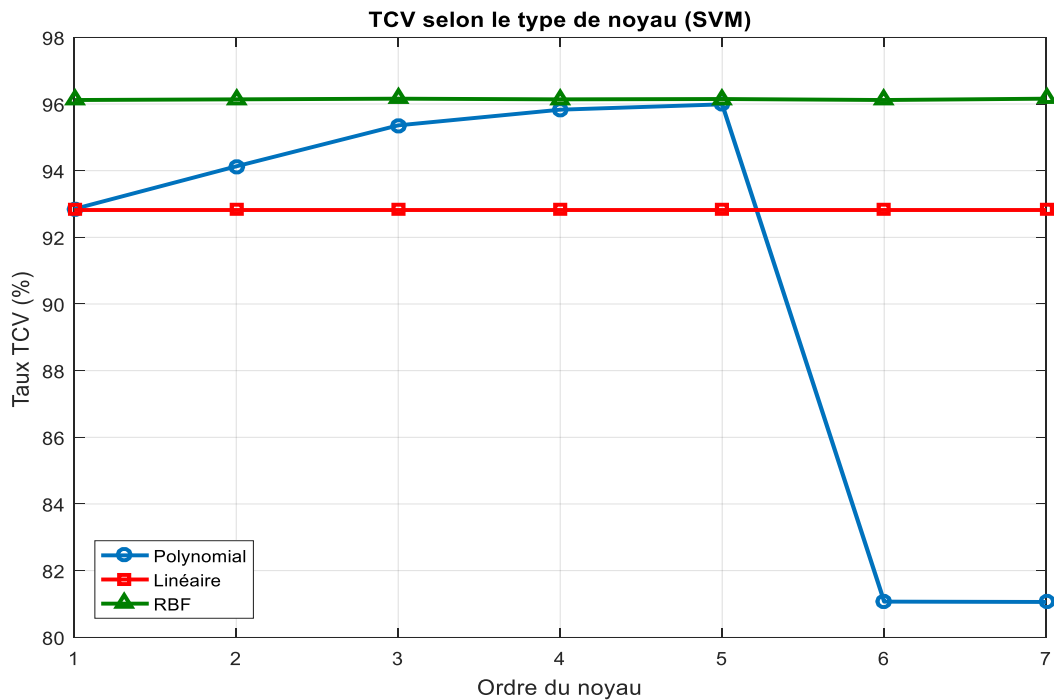


Figure 3.6. Performance TCS selon le type de noyau (SVM).

D'après le Tableau 3.2 et les Figures 3.5 et 3.6 fournis, les observations suivantes peuvent être faites :

- **Noyau RBF** : Ce noyau démontre d'excellentes performances stables. Sa précision TCV est constamment autour de 96 %, et son TCS est de 100 % pour tous les locuteurs. Cela indique que le noyau RBF est le plus robuste et le plus performant parmi ceux testés.
- **Noyau Linéaire** : Le noyau linéaire montre des performances constantes mais inférieures, avec des précisions TCV et TCS autour de 92 %. La présence de résultats manquants pour certains locuteurs suggère que ce noyau pourrait ne pas être adapté pour gérer la complexité de certaines données.
- **Noyau Polynomial** : Le noyau polynomial présente un potentiel mais est très sensible. Bien qu'il puisse atteindre d'excellentes performances (TCS de 100 % pour certains locuteurs), ses performances sont très variables, avec une chute significative pour d'autres locuteurs (autour de 62 %). Cette grande variabilité indique une forte sensibilité aux données ou au locuteur spécifique.

### 3.5.1.3. Système basé sur le modèle 1D-CNN en mode dépendant du locuteur

Dans cette section nous avons présenté et évalué les performances d'un modèle de réseau de neurones convolutif unidimensionnel (1D-CNN) appliqué à la classification de signaux vocaux, en tenant compte des caractéristiques propres à chaque locuteur. Autrement dit, il s'agit de montrer comment le modèle 1D-CNN peut apprendre à reconnaître la MP en s'adaptant spécifiquement à chaque individu, grâce à un entraînement personnalisé par locuteur.

Le tableau 3.3 ci-dessus présente les principaux hyperparamètres sélectionnés pour l'entraînement du modèle 1D-CNN (mode dépendant) . Ces choix ont été faits après plusieurs essais dans des plages de valeurs définies, afin d'optimiser la performance du réseau

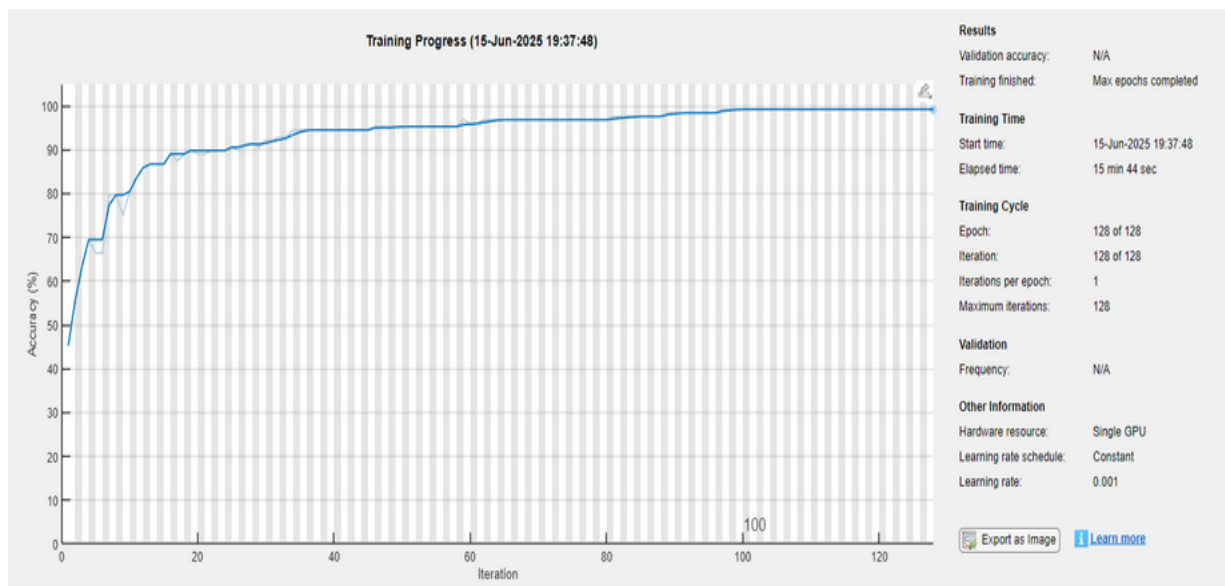
**Tableau 3.3.** Les hyperparamètres clés choisis pour entraîner un réseau de neurones convolutifs unidimensionnel (1D-CNN).

Type de paramètre	Plage de valeurs	Valeur sélectionnée
Nombre de couches de convolution 1D	2–4	3
Nombre de filtres	8–512	64

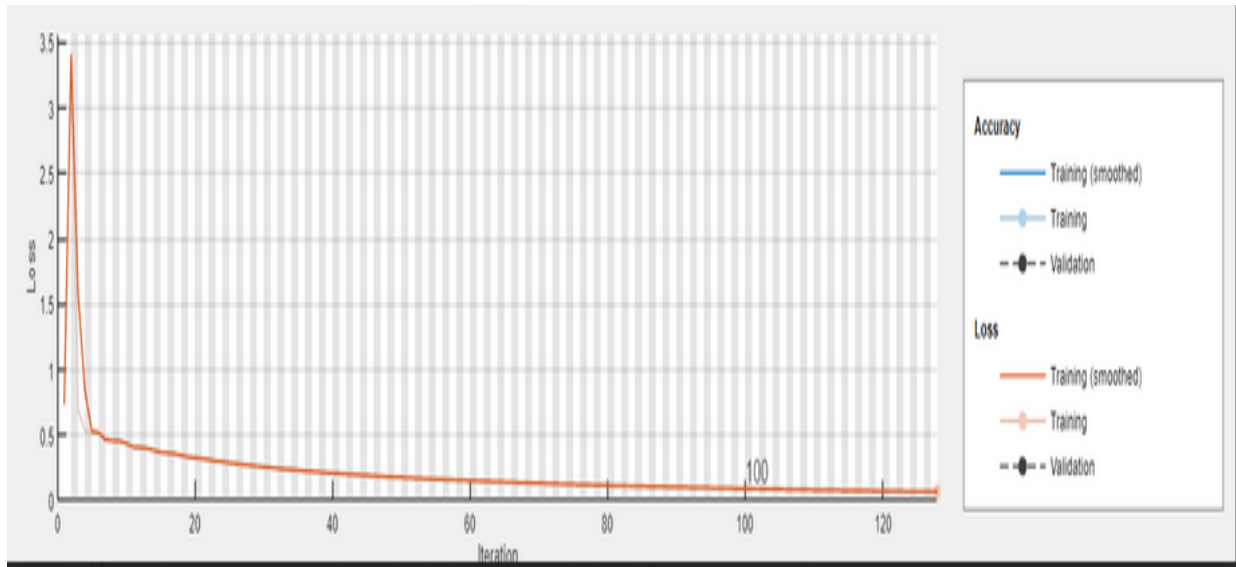
### Chapitre 3 : Conception et implémentation d'un système de diagnostic de la maladie de Parkinson basé sur l'apprentissage profond

Taille du noyau	<b>1×1–7×1</b>	<b>3×1</b>
Taille du lot	<b>8–32</b>	<b>22</b>
Taux d'apprentissage	<b>0.0005–0.01</b>	<b>0.001</b>
Nombre d'époques	<b>50–500</b>	<b>128</b>
Algorithme d'optimisation	<b>Adam – SGDM</b>	<b>Adam</b>

Les figures 3.7 et 3.8 présentes les performances du modèle 1D-CNN en mode dépendant du locuteur.



**Figure 3.7.** Précision d'entraînement au cours des itérations (mode dépendant).



**Figure 3.8.** Perte d'entraînement au cours des itérations (mode dépendant).

A partir des hyperparamètres illustrés sur le tableau 3.3 ainsi que les figures 3.7 et 3.8, nous avons obtenu un taux de reconnaissance de 93,08 % en utilisant un modèle 1D-CNN, ce qui témoigne de son efficacité pour la tâche considérée. Toutefois, ce type de modèle nécessite une base de données volumineuse et représentative pour atteindre de bonnes performances. Dans cette optique, nous prévoyons d'explorer des techniques d'augmentation de données afin d'accroître la diversité des exemples d'apprentissage. Par ailleurs, la fusion avec d'autres modalités, telles que les signaux (EEG, ECG, EMG, IRM, Ecriture ou les images), pourrait constituer une piste prometteuse pour renforcer la robustesse et la précision du système.

### **3.5.2. Système basé sur les classificateurs KNN, SVM et 1D-CNN en mode indépendant du locuteur**

Dans cette section, nous cherchons à identifier la meilleure configuration des classificateurs KNN, SVM et du modèle de réseau de neurones convolutif unidimensionnel (1D-CNN) afin d'obtenir les meilleurs taux de classification des vecteurs (TCV) et le taux de classification des signaux (TCS) pour KNN et SVM et d'obtenir le meilleur taux de classification pour 1D-CNN en mode indépendant du locuteur.

### 3.5.2.1. Système basé sur les classificateurs KNN en mode indépendant du locuteur

L'objectif est d'optimiser le classificateur KNN pour le mode indépendant du locuteur, en sélectionnant le nombre de voisins K et la distance la plus adaptée (Euclidean, City Block, Cosine ou Correlation), afin d'améliorer les performances de classification (TCV et TCS) du système de diagnostic.

Le tableau suivant présente les taux de classification des vecteurs (TCV) ainsi que les taux de classification des signaux par vote majoritaire (TCS) obtenus à l'aide du classificateur K plus proches voisins (KNN), en testant différentes valeurs de K (1 à 25) et divers types de distances : Euclidean, Cityblock, Cosine, et Correlation.

**Tableau 3.4.** Performance du classificateur KNN selon le nombre de voisins et le type de distance (mode indépendant du locuteur).

Distance	Euclidean		Cityblock		Cosine		Correlation	
	TCV (%)	TCS (%)	TCV (%)	TCS (%)	TCV (%)	TCS (%)	TCV (%)	TCS (%)
1	77.13	79.14	76.80	79.57	73.13	76.60	71.99	74.89
2	77.68	80.85	77.60	80.85	73.72	77.44	72.52	75.74
3	77.21	79.14	77.17	79.57	73.08	75.32	71.89	74.04
4	77.61	78.72	77.59	80.85	73.49	77.87	72.18	74.89
5	77.25	77.87	77.25	80.00	73.03	75.74	71.77	74.04
6	77.50	78.29	77.56	80.85	73.33	76.17	71.98	74.47
7	77.27	77.87	77.26	80.42	72.95	75.32	71.65	73.62
8	77.44	77.44	77.54	80.85	73.21	76.17	71.84	74.04
9	77.24	77.87	77.27	79.57	72.94	75.32	71.55	73.62
10	77.35	77.87	77.47	80.85	73.12	76.17	71.69	74.04
11	77.23	77.87	77.47	79.14	72.87	75.32	71.48	73.61
12	77.35	77.87	77.42	80.85	73.02	75.74	71.60	74.04
13	77.21	77.44	77.23	80.00	72.80	75.32	71.39	73.61
14	77.30	77.83	77.40	80.85	72.93	75.74	71.46	74.04
15	77.16	77.44	77.20	80.42	72.70	75.32	71.28	73.19
16	77.24	77.44	77.37	80.43	72.84	75.74	71.41	74.04
17	77.12	77.44	77.20	80.43	72.65	75.32	71.22	73.19

18	77.20	77.44	77.33	80.43	72.76	76.60	71.32	73.19
19	77.05	77.87	77.18	80.43	72.57	74.89	71.13	72.76
20	77.12	77.44	77.27	80.43	72.67	74.89	71.23	72.77
21	77.00	77.44	77.13	80.43	72.48	75.32	71.08	72.77
22	77.10	77.44	77.25	80.43	72.60	75.32	71.17	72.34
23	76.99	77.44	77.11	80.43	72.42	75.32	71.03	72.77
24	77.02	77.44	77.21	80.00	72.55	74.89	71.11	72.77
25	76.93	77.02	77.09	80.43	72.39	74.89	70.95	72.34

Ces résultats sont illustres dans les figures 3.9 et 3.10.

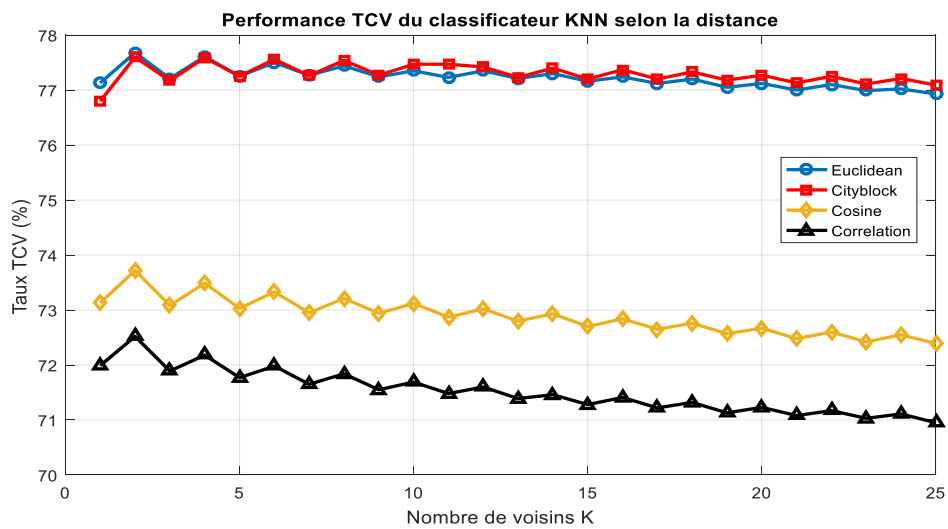


Figure 3.9. Performance TCV du classificateur KNN selon la distance.

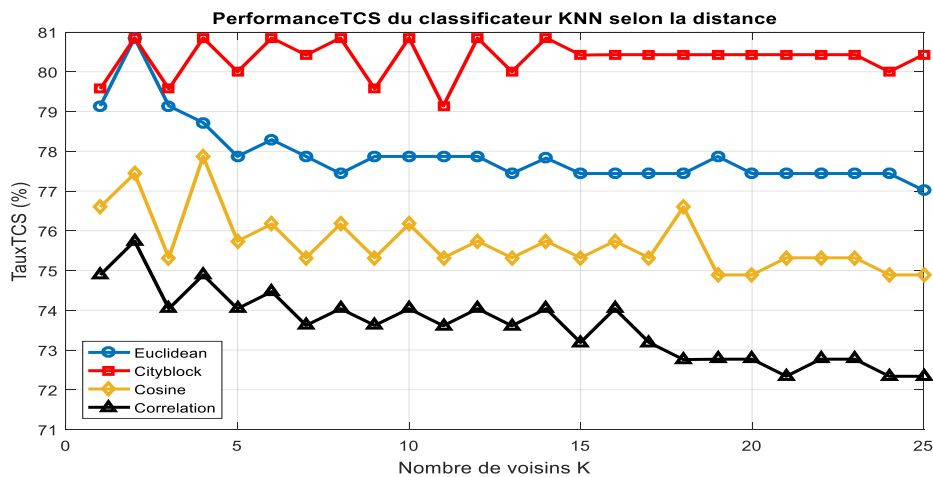


Figure 3.10. Performance TCS du classificateur KNN selon la distance.

A partir des résultats illustrés sur le tableau 3.4 ainsi que les figures 3.9 et 3.10, nous pouvons remarquer les points suivants :

- Les performances du classificateur KNN sont fortement dépendantes du choix de la métrique de distance.
- La **distance Cityblock** a clairement dominé en termes de TCS, atteignant des pics allant jusqu'à 80,85 %, et a également fourni d'excellents résultats TCV (~77,61 %), se classant parmi les meilleures.
- La **distance Euclidienne** a affiché des performances très comparables à Cityblock pour le TCV et de bonnes performances pour le TCS, bien que légèrement inférieures à Cityblock.
- Les **distances Cosine et Correlation** ont systématiquement montré des performances inférieures pour les deux métriques par rapport aux distances Euclidienne et Cityblock.
- Pour la plupart des métriques, les performances tendent à se stabiliser ou à diminuer légèrement avec l'augmentation du nombre de voisins K après des pics initiaux pour de petites valeurs de K.

### 3.5.2.2. Système basé sur les classificateurs SVM en mode indépendant du locuteur

Cette section présente les performances du classificateur SVM appliqué en mode indépendant du locuteur, où le modèle est entraîné et testé sur des données issues de locuteurs différents. Le tableau 3.5 présente ces performances.

**Tableau 3.5.** Performance du classificateur SVM selon le type des noyaux (mode indépendant du locuteur).

Order	Accuracy (%)					
	Polynomial		Linear		Rbf	
	TCV	TCS	TCV	TCS	TCV	TCS
1	<b>80.81</b>	<b>82.98</b>	<b>80.88</b>	<b>3.40</b>	<b>86.95</b>	<b>91.49</b>
2	<b>81.34</b>	<b>84.26</b>			<b>87.19</b>	<b>91.49</b>
3	<b>84.12</b>	<b>87.66</b>			<b>87.12</b>	<b>91.49</b>
4	<b>85.60</b>	<b>88.09</b>			<b>87.14</b>	<b>91.91</b>
5	<b>87.48</b>	<b>90.64</b>			<b>87.18</b>	<b>91.49</b>

6	51.59	53.19			87.05	91.49
7	51.31	53.19			87.10	91.49

Les figures 3.11 et 3.12 ci-dessous présentent les résultats du SVM en mode indépendant du locuteur, en fonction des paramètres utilisés.

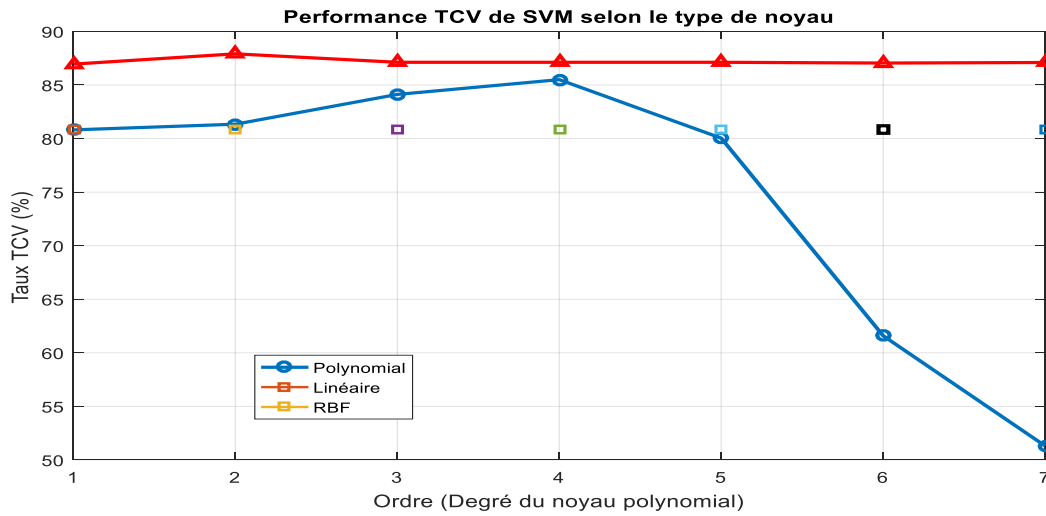


Figure 3.11. Performance TCV de SVM selon le type de noyau.

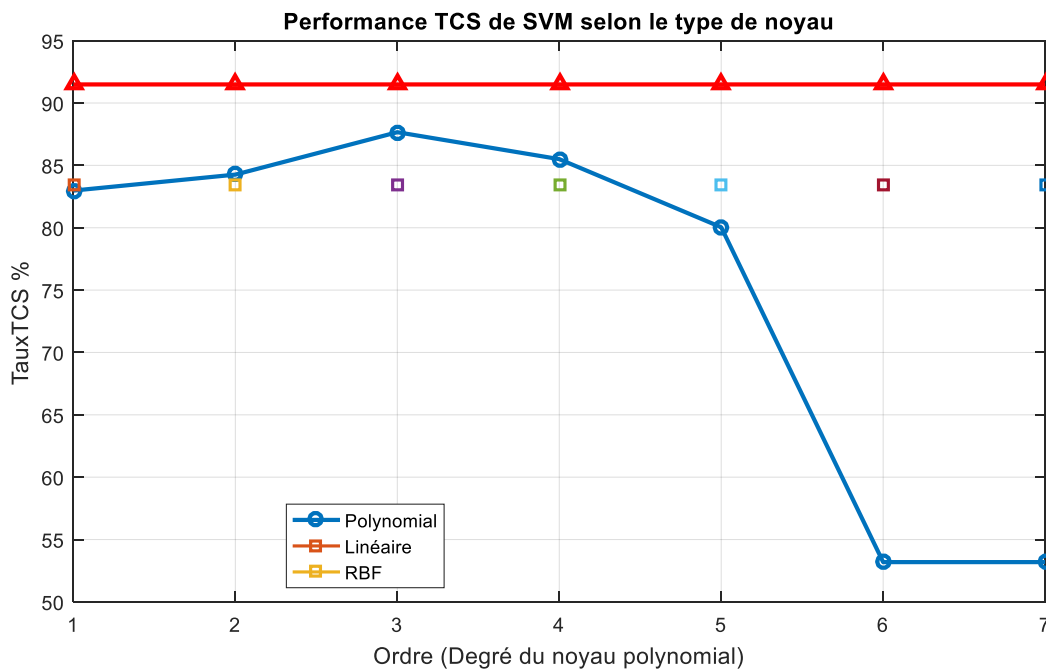


Figure 3.12. Performance TCS de SVM selon le type de noyau

À partir des résultats illustrés sur le tableau 3.6 ainsi que les figures 3.11 et 3.12, nous pouvons remarquer les points suivants :

- **Noyau Polynomial :**

- **Performances très variables :** Ce noyau démontre une grande sensibilité aux données, se traduisant par une variabilité significative des performances.
- **Meilleures valeurs :** Il peut atteindre d'excellentes précisions, notamment pour l'ordre 5, avec un TCV de 87,48 % et un TCS de 90,64 %.
- **Très faibles performances :** Cependant, pour les ordres 6 et 7, les performances chutent drastiquement à des TCV d'environ 51,5 % et des TCS d'environ 53,19 %. Cette forte variabilité est clairement visible sur les figures 3.10 et 3.11, où la courbe "Polynomial" présente une nette chute après l'ordre 5, indiquant une sensibilité élevée à la variabilité inter-locuteur ou à la complexité croissante.

- **Noyau Linéaire :**

- **Performances stables et bonnes :** Le noyau Linéaire affiche des performances stables et respectables avec un TCV de 80,88 % et un TCS de 83,40 %. Ces résultats sont cohérents et indiquent une capacité à bien généraliser sans la complexité excessive des autres noyaux. Les figures 3.10 et 3.11 montrent une ligne plate pour le linéaire, confirmant cette constance.

- **Noyau RBF :**

- **Très stable et cohérent :** Le noyau RBF se distingue par sa robustesse et sa grande régularité de performance. Tous les locuteurs (ordres) montrent un TCS de 91,49 %, à l'exception du locuteur 4 qui atteint un léger pic à 91,91 %.
- **Excellente régularité du TCV :** Les taux TCV varient très peu, de 86,95 % à 87,19 %, ce qui confirme une excellente régularité et une faible sensibilité aux variations des données.
- Les figures 3.10 et 3.11 illustrent parfaitement cette stabilité, la courbe "RBF" restant quasi-constante au plus haut niveau de performance pour l'ensemble des

ordres. Le noyau RBF semble donc être le plus adapté et le plus performant pour le SVM dans ce contexte spécifique.

### 3.5.2.3. Système basé sur le modèle 1D-CNN en mode indépendant du locuteur

Dans cette section, nous avons évalué les performances du modèle 1D-CNN en mode indépendant du locuteur, où l'entraînement et le test sont réalisés sur des locuteurs différents. Le tableau 3.6 montre ces performances.

**Tableau 3.6.** Les hyperparamètres clés choisis pour entraîner un réseau de neurones convolutifs unidimensionnel (CNN 1D).

Type de paramètre	Plage de valeurs	Valeur sélectionné
Nombre de couches de convolution 1D	<b>2–4</b>	<b>3</b>
Nombre de filtres	<b>8–512</b>	<b>48</b>
Taille du noyau	<b>1×1–7×1</b>	<b>3×1</b>
Taille du lot	<b>8–32</b>	<b>22</b>
Taux d'apprentissage	<b>0.0005–0.01</b>	<b>0.001</b>
Nombre d'époques	<b>50–500</b>	<b>128</b>
Algorithme d'optimisation	<b>Adam – SGDM</b>	<b>Adam</b>

Les figures 3.13 et 3.14 présentent les performances du modèle 1D-CNN en mode dépendant du locuteur.

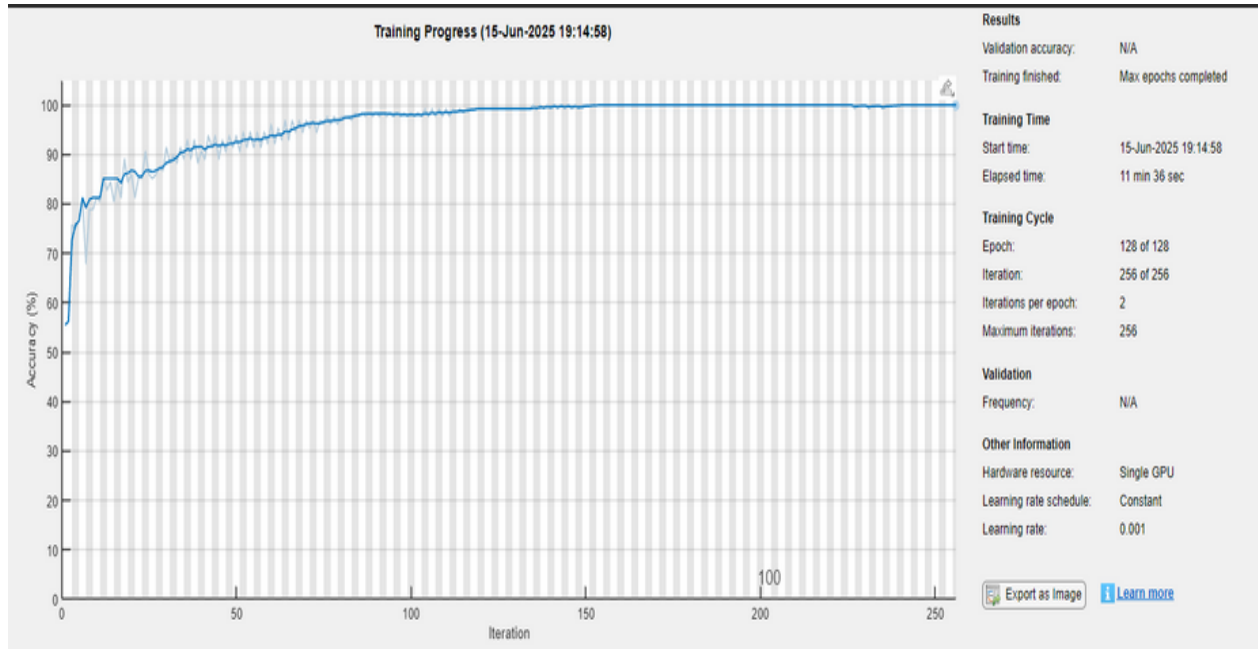


Figure 3.13. Précision d'entraînement au cours des itérations (mode indépendant).

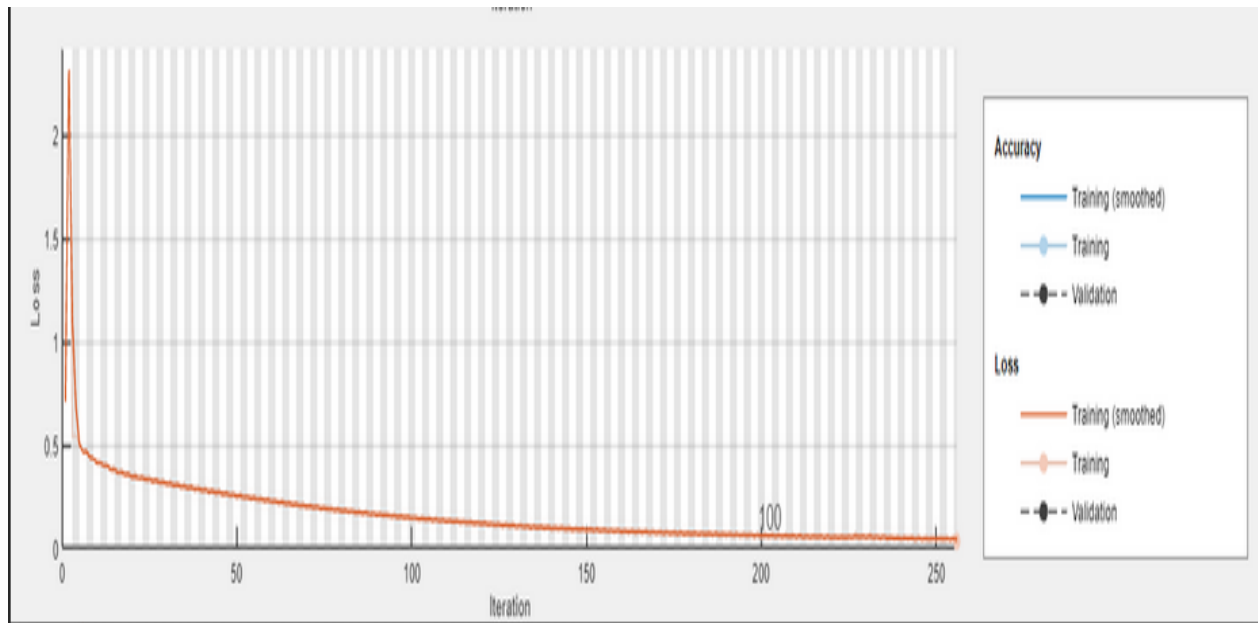


Figure 3.14. Perte d'entraînement au cours des itérations (mode indépendant).

A partir des hyperparamètres illustrés sur le tableau 3.6 ainsi que les figures 3.13 et 3.14, nous avons obtenu un taux de reconnaissance de 83,82 % en utilisant un modèle 1D-CNN en mode

indépendant, ce qui montre un potentiel intéressant malgré des performances légèrement inférieures à celles obtenues en mode dépendant. Ce résultat met en évidence la capacité du réseau à généraliser sur des locuteurs non vus durant l'entraînement, mais souligne également les limites liées à la variabilité interindividuelle. Pour améliorer ces performances, il serait pertinent d'explorer des techniques d'augmentation de données, permettant d'enrichir la diversité des échantillons vocaux. En outre, la fusion multimodale avec d'autres types de signaux biomédicaux (tels que EEG, ECG, EMG, IRM, ou encore l'écriture manuscrite et les images) représente une voie prometteuse pour renforcer la robustesse et la précision du système de diagnostic automatique.

### **3.6. Etude comparatives des performances des classificateurs KNN, SVM et la méthode 1D-CNN en mode dépendant et indépendant de locuteur**

Le tableau Tab.3.7 illustre les résultats comparatifs des différents classificateurs KNN, SVM et 1D-CNN en terme du taux TCS en modes dépendant et indépendant du locuteur.

**Tableau 3.7.** Les hyperparamètres clés choisis pour entraîner un réseau de neurones convolutifs unidimensionnel (CNN 1D)

<b>Critère</b>	<b>KNN</b>	<b>SVM (Rbf)</b>	<b>1D-CNN</b>
<b>Type</b>	Maching learning	Maching learning	Deep learning
<b>TCS (%)</b> <b>Mode dépendant</b>	99.54 %	100 %	93.08%
<b>TCS (%)</b> <b>Mode indépendant</b>	80.85 %	91.91%	83.82%
<b>Temps de calcul</b>	Faible à moyen	Moyen	Élevé

Parmi ces résultats, il est possible de tirer une comparaison claire entre les trois méthodes, en mettant en évidence les points forts et les limites de chacune selon le mode de classification (dépendant ou indépendant).

- SVM (RBF) atteint 100 %, ce qui montre une excellente capacité à s'adapter aux spécificités individuelles lorsqu'il est entraîné et testé sur le même locuteur.
- KNN suit de près avec 99.54 %.

- 1D-CNN, bien qu'étant une méthode avancée, obtient 93.08 %, ce qui est légèrement inférieur. Cela peut être dû au fait qu'un entraînement plus long ou un affinement des hyperparamètres est nécessaire.

En mode indépendant

- Ici, SVM conserve sa supériorité avec 91.91 %, montrant une bonne capacité de généralisation à de nouveaux locuteurs.
- 1D-CNN fait mieux que KNN avec 83.82 %, grâce à sa capacité à capturer des modèles temporels complexes.
- KNN chute à 80.85 %, ce qui reflète une difficulté à généraliser lorsqu'il est confronté à des voix qu'il n'a jamais vues.

### **3.7. Conclusion**

Dans ce travail, nous avons comparé plusieurs classificateurs – KNN, SVM et 1D-CNN pour détecter la MP à partir de la voix. Les résultats montrent que le SVM, notamment avec un noyau RBF, donne d'excellentes performances, atteignant jusqu'à 100 % de précision en test. Le modèle 1D-CNN, quant à lui, a obtenu 93.08 % en mode dépendant et 83.82 % en mode indépendant, ce qui confirme son efficacité malgré une sensibilité à la variabilité des locuteurs. Pour améliorer encore les performances, des approches comme l'augmentation de données ou la fusion avec d'autres signaux biomédicaux sont envisagées.

## Conclusion générale

Dans ce mémoire, nous avons développé un système de diagnostic automatique de la MP à partir de signaux vocaux, en combinant des techniques d'analyse du signal avec des méthodes d'intelligence artificielle. L'objectif principal était d'évaluer l'efficacité de différents classificateurs pour distinguer les voix saines des voix parkinsoniennes à partir des caractéristiques acoustiques MFCC.

Nous avons d'abord exploré des approches classiques de machine learning telles que K-Nearest Neighbors (KNN) et Support Vector Machines (SVM). Ces méthodes ont montré de bonnes performances, en particulier le SVM avec noyau RBF, qui a atteint une précision élevée en modes dépendant et indépendant, démontrant ainsi sa capacité à modéliser des frontières de décision complexes.

Nous avons ensuite mis en œuvre un modèle d'apprentissage profond basé sur un réseau de neurones convolutifs 1D (1D-CNN), appliqué directement aux vecteurs MFCC. Ce modèle a fourni de bons résultats, avec une précision de 93,08 % en mode dépendant et 83,82 % en mode indépendant. Néanmoins, ces résultats montrent une précision faible de cet algorithme par rapport aux algorithmes de machine learning qui montrent plus de robustesse en mode indépendant du locuteur. Ceci est due probablement à cause de la limitation de la base d'apprentissage.

Dans l'ensemble, les résultats obtenus confirment que l'utilisation des MFCC combinée à des classificateurs performants constitue une approche prometteuse pour un diagnostic non invasif de la MP. Pour améliorer davantage la robustesse du système, des perspectives intéressantes incluent l'utilisation de techniques d'augmentation de données, d'adaptation au locuteur, ou encore la fusion multimodale avec d'autres signaux biomédicaux (EEG, EMG, écriture manuscrite, etc.).

## Bibliography

- [1] Marion, Berthon,, «“Maladies neurologiques: ce qu'il faut savoir”,» 28/06/2021. [En ligne].
- [2] S. L. KOWAL, T. M. DALL, . R. CHAKRABARTI et et al., «"The current and projected economic burden of Parkinson's disease in the United States",» *Movement Disorders*, vol. 28, n° %13, pp. 311-318, 2013.
- [3] P. J, S. J.R, G. V.D, A.K et M. H. V. F.W, «Parkinson’s disease is associated with dysregulations of a dopamine-modulated gene network relevant to sleep and affective neurobehaviors in the striatum,» 2019.
- [4] W. W. J., S. L. M et L. A. E, «Parkinson's disease: A complete guide for patients and families“. Baltimore : The Johns Hopkins University Press,» 2013.
- [5] D. A et . B. and S, «Classification automatique de la maladie de Parkinson à partir de la voix,» Université de Bordj Bou Arreridj Faculty of Mathematics and Computer Science, Bordj Bou Arreridj, 2024.
- [6] G. W. P et et al, «“A common LRRK2 mutation in idiopathic Parkinson's disease”,» *The Lancet*, vol. Vol.365, p. pp. 415–416, 2005.
- [7] K. L et . R. and H, «“Pathogenesis of Parkinson disease—the gut–brain axis and environmental factors”,» *Nature Reviews Neurology*, vol. vol. 11, n° %1no. 11, p. pp. 625–636, 2015.

- [8] S. R. , P. J , S. H, P. A et R. S. and P, «“Slowing of EEG in Parkinson's disease”,» *Electroencephalography and Clinical Neurophysiology* , vol. vol. 79 , p. pp. 159–165, Sept 1991 .
- [9] C. J, «“Use of magnetoencephalography (MEG) to study functional brain networks in neurodegenerative disorders”,» *Journal of the Neurological Sciences*, vol. vol. 289, n° %1no. 1–2, p. pp. 128–134, 2010.
- [10] H. T. H, P. V, . K. J. T, H. H. V, S. K. A et M. and V. V, «“Ambulatory ECG and analysis of heart rate variability in Parkinson’s disease”,» *Journal of Neurology, Neurosurgery & Psychiatry*, vol. vol. 70, n° %1no. 3, p. pp. 305–310, Mar 2001.
- [11] P. N, G. C, G.-L. M. D et . L. and S, «“A review of the use of magnetic resonance imaging in Parkinson’s disease”,» *Therapeutic Advances in Neurological Disorders* , vol. vol. 7 , n° %1no. 4 , p. pp. 206–220, 2014.
- [12] D. P et et al, «“Prediction potential of different handwriting tasks for diagnosis of Parkinsons”,» chez *Proc. E-Health and Bioengineering Conference (EHB)* , 2013.
- [13] I. Aouraghe, «" Développement de méthodes de l’intelligence artificielle et de l’apprentissage automatique pour la détection précoce et la classification des pathologies neurodégénératives basé sur l’écriture manuscrite Arabe en ligne sur tablette graphique",» 2020.
- [14] A. Le Roux et A. et Thomas, «"IMPACT D’UNE RÉÉDUCATION VOCALE INTENSIVE SUR LA DYSPROSODIE PARKINSONNIENNE",» Université Claude bernard Lyon 1, 30 Juin 2011.
- [15] L. Jeancolas, «"Détection précoce de la maladie de Parkinson par l’analyse de la voix et corrélations avec la neuroimagerie",» Paris, 2019.

- [16] M. Grueber, «"Effets de la stimulation sous-thalamique bilatérale sur la voix et la parole de patients parkinsoniens",» Lausanne, 2011.
- [17] X. Xiong, «Robust speech features and acoustic models for speech recognition,» 2009.
- [18] T. Drissi, Belhoussine et et al, «"Diagnosis of Parkinson's disease based on wavelet transform and mel frequency cepstral coefficients",» *International Journal of Advanced Computer Science*, vol. vol.10, n° %1n.3, p. 125–132, 2019.
- [19] B. Hichem et et al, «"Surface EMG Signal Classification for Parkinson's Disease using WCC Descriptor and ANN Classifier.",» 2021.
- [20] H. F. MEROUANI et et al, «L'Apprentissage l'Interprétation.».
- [21] M. M. Giulia, «Intelligence artificielle et maladies neurologiques : aider le diagnostic et améliorer la,» 2022..
- [22] W. BOUSSOUALIM et . S. & IDIR, «"Classification des IRM cérébrales pathologiques basée sur les techniques d'apprentissage profond",» Bordj Bou Arréridj, 2023.
- [23] S. KHESSIBA, «"Stratégies d'optimisation des hyper-paramètres de réseaux de neurones appliqués aux signaux temporels biomédicaux",» 2024.
- [24] [En ligne]. Available: <https://www.ibm.com/fr-fr/topics/supervised-learning>.
- [25] H. Bengacemi, , A. Hacine-Gharbi, P. Ravier, K. Abed-Meraim et O. Buttelli, «Surface EMG Signal Classification for Parkinson's Disease using WCC Descriptor and ANN Classifier,» chez *Proceedings of the 10th International Conference on Pattern Recognition Applications and Methods (ICPRAM)*, Février 2021.

- [26] C. CHAKOUR, M. Z. BASSA et . M. A. BENLAMNOUAR, «Méthodes d'intelligence artificielle appliquées au diagnostic de défauts».
- [27] N. DEGDEG, «"Application de l'apprentissage automatique pour un problème de classification des machines de l'entreprise Mega-Papiers",» 2022.
- [28] S. MANSOUR et Z. TABTI, «Sélection de variables pour la reconnaissance de la maladie de Parkinson,» 2016.
- [29] O. Zammit, . X. Descombes et J. Zerubia, «"Apprentissage non supervisé des SVM par un algorithme des K-moyennes entropique pour la détection de zones brûlées",» 2007.
- [30] F. Ghazali, A. Hacine-Gharbi, K. Rouabah et P. & Ravier, «Performance Evaluation of the Electrical Appliances Identification System Using the PLAID Database in Independent Mode of House.,» chez *ICPRAM*, 2024, February.
- [31] N. Ben Chaabane, «Analyse de signaux physiologiques pour la prédiction de l'évolution pathologique par intelligence artificielle : Application à la maladie de Parkinson et à l'analyse quantifiée de la marche,» 2023.
- [32] T. Claude, *Les réseaux de neurones artificiels : introduction au connexionnisme*, EC2, Éd., 1992.
- [33] M. Clergue, *Réseaux de neurones artificiels*, 2015.
- [34] G. Gelly, «Réseaux de neurones récurrents pour le traitement automatique de la parole,» 2017.

- [35] Z. Aamouche et Marouane, «: Classification des battements cardiaques en utilisant les réseaux de neurones profonds,» 2019.
- [36] D. Chekari et . N. Boumehti, «Réalisation d'un système d'aide au diagnostic des images mammographiques basé sur le Cloud,» Tiaret.
- [37] A. A. Toye et S. Kompalli, «Comparative study of speech analysis methods to predict Parkinson's disease,» 2021.
- [38] I. Ahmed et et al, «Classification of Parkinson disease based on patient's voice signal using machine learning,» vol. 32, n° %12, p. 705, 2022.
- [39] R. Alshammri et et al, «Machine learning approaches to identify Parkinson's disease using voice signal features,» vol. 6, 2023.
- [40] M. Meral, . F. Ozbilgin et F. Durmus, «Fine-Tuned Machine Learning Classifiers for Diagnosing Parkinson's Disease Using Vocal Characteristics: A Comparative Analysis,» vol. 15, n° %15, 2025.
- [41] I. Karabayir et et al, «Gradient boosting for Parkinson's disease diagnosis from voice recordings,» vol. 20, p. 1–7, 2020.
- [42] C. Quan et et al., «End-to-end deep learning approach for Parkinson's disease detection from speech signals,» vol. 42, n° %12, p. 556–574, 2022.
- [43] G. Dimauro et G. and Francesco , «Italian parkinson's voice and speech.",» 2019). [En ligne]. Available: [10.21227/aw6b-tg17](https://doi.org/10.21227/aw6b-tg17) .

