

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 mathématiques et informatique

Département informatique

Mémoire

Présenté pour obtenir

LE DIPLOME DE MASTER

FILIERE : INFORMATIQUE

Spécialité : Technologies de l'information et de la communication

THEME :

*Classification automatique de la maladie
de Parkinson à partir de la voix*

Présenté Par :

- Dahili Ahlem
- Benmessahel Samira

Soutenu le : 19/06/2024

Devant le jury composé de :

Président : Mr Charikhi Mourad
Examineur : Mr Saifi Abdelhamid
Encadreur : Mr Belazoug Mouhoub

Année Universitaire 2023/2024



Dédicace

A l'homme de ma vie, mon exemple éternel, mon soutien moral et source de joie et de bonheur, celui qui s'est toujours sacrifié pour me Voir réussir, que dieu te garde dans son vaste paradis, à toi mon mari

A la lumière de mes jours, la source de mes efforts, la flamme de mon cœur, ma vie et mon bonheur ; maman que j'adore.

mes chères frères Adel et Hocine et chères sœurs Warda et Djahida

Mes chers enfants yahya et Aya

Aux personnes qui m'ont toujours aidé et encouragé, qui étaient toujours à mes côtés, et qui m'ont accompagnaient durant mon chemin D'études supérieures, mes aimables amis, tous mes collègues d'étude.

A mon binôme Ahlem et toute la famille Benmessahel. Et à tous ceux qui ont contribué de près ou de loin pour Que ce projet soit possible, Tous ceux que j'aime dans le monde. Je vous Dis merci.

SAMIRA





Dédicace

A l'homme de ma vie "papa", mon exemple éternel, mon soutien moral et source de joie et de bonheur, celui qui s'est toujours sacrifié pour me voir réussir, Que Dieu ait pitié de toi mon père, à toi mon père.

A la lumière de mes jours, la source de mes efforts, la flamme de mon cœur, ma vie et mon bonheur ; maman que j'adore.

Mes très chères frères Hamza, Aymen, Wael.

Aux personnes qui m'ont toujours aidé et encouragé, qui étaient toujours à mes côtés, et qui m'ont accompagnaient durant mon chemin D'études supérieures, mes aimables amis, tous mes collègues d'étude.

A mon binôme Samira et toute la famille Dahili. Et à tous ceux qui ont contribué de près ou de loin pour que ce projet soit possible, Tous ceux que j'aime dans le monde. Je vous

Dis merci.

AHLEM



Remerciement

Nous tenons à exprimer nos sincères remerciements à tous les professeurs qui nous ont enseigné et qui, par leur expertise, nous ont accompagnés dans nos démarches académiques.

*En second lieu, nous tenons à exprimer notre gratitude particulièrement au **Dr. Belazoug Mouhoub** pour ses conseils, sa confiance, sa patience et son encadrement tout au long de notre thèse, pour les nombreux conseils qu'il nous a prodigués, ainsi que pour le temps qu'il a passé avec nous pour assurer la réussite de cette thèse.*

Nous remercions également l'équipe académique de notre faculté de Mathématiques et d'Informatique pour avoir répondu à nos questions tout au long de ces cinq années.

Nos sincères remerciements vont également aux Membres du Jury pour l'intérêt qu'ils ont porté à nos recherches, pour avoir accepté d'évaluer notre travail et pour l'avoir enrichi de leurs suggestions.

Résumé

L'analyse de la parole offre une approche prometteuse pour le diagnostic précoce et automatisé de la maladie de Parkinson. Cette méthode non invasive et peu coûteuse s'appuie sur les altérations vocales caractéristiques de la maladie, présentes dès ses premiers stades, pour identifier les patients. Des systèmes automatisés basés sur l'intelligence artificielle permettent d'analyser ces altérations vocales et de discriminer efficacement les patients atteints de la maladie de Parkinson des sujets sains.

Malgré des défis tels que la variabilité de la voix et la présence de bruit de fond, l'analyse de la parole présente un fort potentiel pour améliorer le diagnostic et la prise en charge de la maladie de Parkinson. Des recherches continues visent à perfectionner cette technologie et à en faire un outil précieux pour améliorer la qualité de vie des patients.

Abstract

Speech analysis is a promising approach for early and automated diagnosis of Parkinson's disease. This non-invasive and inexpensive method relies on the characteristic voice changes of the disease, present from the early stages, to identify patients. Automated systems based on artificial intelligence can analyze these voice changes and effectively discriminate Parkinson's disease patients from healthy subjects.

Despite challenges such as voice variability and background noise, speech analysis has great potential to improve the diagnosis and management of Parkinson's disease. Ongoing research aims to refine this technology and make it a valuable tool for improving the quality of life for patients.

المخلص

تحليل الكلام يقدم نهجًا واعدًا للكشف المبكر والتلقائي عن مرض باركنسون. تعتمد هذه الطريقة غير الغازية وغير المكلفة على التغيرات الصوتية المميزة للمرض، والتي توجد في مراحله المبكرة، لتحديد المرضى. يمكن لأنظمة الذكاء الاصطناعي الآلية تحليل هذه التغيرات الصوتية وتمييز مرضى باركنسون عن الأشخاص الأصحاء بشكل فعال.

على الرغم من التحديات مثل تنوع الأصوات ووجود الضوضاء الخلفية، فإن تحليل الكلام لديه إمكانات كبيرة لتحسين تشخيص مرض باركنسون وإدارته. تهدف الأبحاث الجارية إلى صقل هذه التكنولوجيا وجعلها أداة قيمة لتحسين نوعية حياة المرضى.

Liste des acronymes et abréviations :

MP : Maladie de parkinson.

SCP : Simulation créable profond

COMT : catéchol – o – méthyltransférase.

MMSE : état mental min – mental

IA : Intelligence artificiel.

ML : Machine learning.

DL : Deep learning.

KNN : K plus proches voisins .

SVP : Support vector machine.

GMM : Modèle mélange gaussien.

Table des matières

Dédicace	2
Remerciements	4
Résumé	5
Liste des acronymes et abréviations	6
Table des matières.....	7
Liste des figures.....	9
Liste des tableaux.....	9
Introduction générale.....	11
Chapitre 01 : la maladie de parkinson	
1.Introduction.....	14
2.La maladie parkinson.....	14
3.Pathophysiologie.....	15
4.La progression de la maladie parkinson.....	16
5.La symptomatologie.....	18
6.La diagnostic clinique	19
7.Traitement	20
7.1Médicaments pour les symptômes moteurs.....	21
7.2Traitement chirurgicale	23
8.La maladie de parkinson et la parole	26
9.Conclusion	27

Chapitre 02 : Apprentissage Automatique

1.Introduction.....29

2.Apprentissage automatique31

3.L’impact de l’apprentissage automatique sur le domaine de la santé31

4.Différents types d’apprentissage32

5.Les algorithmes d’apprentissage automatique.....38

6.Apprentissage profond.....43

7.Conclusion.....44

Chapitre 03 : la méthodologie

1.Introduction.....46

2.Taches46

3.Présentations de dataset46

4.Les mesures de performances utilisées50

5.Conclusion.....51

Chapitre 04 : expérimentations et résultats

1.Introduction53

2.Présentations des outils utilisées55

3.Expérimentation et discussion58

4.Comparaisons entre les résultats62

5.Conclusion.....62

Conclusion générale.....64

Références.....65

Liste des figures

Figure 1: Location of substance noire et la différence entre normal et la maladie parkinson.....	15
Figure2 : Illustration de la stadification de braak dans la maladie de parkinson.....	17
Figure3: Deep brain stimulation	24
Figure 4: Stem cell transplantations	25
Figure 5 : La relation entre IA, ML, DL	29
Figure 6 : Apprentissage supervisé	32
Figure 7 : Apprentissage non supervisé	33
Figure 8 : Apprentissage supervisé vs non supervisé.....	34
Figure 9 : Classification vs régression	34
Figure 10 : Apprentissage par renforcement	36
Figure 11 : Classification a deux classes avec séparatrice linéaire.....	38
Figure 12 : Classification 2 classes avec une séparatrice non linéaire.....	38
Figure 13 : Régression logistique	39
Figure 14 : Modèle de mélange gaussien	41
Figure 15 : Le modèle proposé	53
Figure 16 : Logo python.....	54
Figure 17 : Visual studio code.....	55
Figure 18 : Interface Visual code.....	55
Figure 19 : Séparation des données d'entrainements et de test.....	57
Figure 20 : Comparaison de résultats des algorithmes d'apprentissage régression logistique	58
Figure 21 : Comparaison de résultats des algorithmes d'apprentissage SVM.....	59
Figure 22 : Comparaison de résultats des algorithmes d'apprentissage GMM	60

Liste des Tableau

Tableau 1 : comparaison entre les résultats des algorithmes d'apprentissage	61
--	----

Introduction générale

Introduction :

La maladie de Parkinson (MP) est un trouble neurodégénératif complexe qui a été décrit pour la première fois en 1817 par le médecin James Parkinson. Il est caractérisé par des symptômes moteurs et non moteurs qui s'aggravent avec le temps. Si le diagnostic clinique est clair aux stades avancés, il devient difficile aux stades précoces lorsque les symptômes sont subtils. Cela peut conduire à des retards dans le diagnostic et la nécessité de visites cliniques fréquentes pour surveiller la progression de la maladie. Cette difficulté de détection précoce et de suivi de la progression de la MP a motivé la recherche de nouveaux outils d'évaluation informatisés ou d'aide à la décision pour un diagnostic précoce et une meilleure prédiction de l'évolution de la maladie. Des études ont montré que les déficiences vocales, les changements dans l'écriture manuscrite et les mouvements oculaires saccariques peuvent être des indicateurs précoces de la maladie. Ces déficiences pourraient jouer un rôle crucial dans la détection précoce et la surveillance de la progression de la maladie une fois diagnostiquée.

La maladie de Parkinson représente un défi majeur dans le domaine de la santé, étant un trouble neurodégénératif complexe qui affecte la qualité de vie des patients. Avec des symptômes moteurs et non moteurs qui évoluent avec le temps, la détection précoce et le suivi de la progression de la maladie deviennent essentiels pour une gestion efficace de la condition. Cependant, le diagnostic précoce de la maladie de Parkinson demeure difficile, en particulier aux stades initiaux lorsque les symptômes sont subtils, ce qui peut entraîner des retards dans la prise en charge clinique et des complications pour les patients.

Afin de pallier cette difficulté, il est impératif de développer des outils de détection précoce efficaces, économiques, facilement accessibles et non invasifs. Dans cette optique, l'utilisation de signaux vocaux comme indicateur potentiel de la maladie de Parkinson offre une voie prometteuse. Des études ont déjà suggéré que les altérations vocales pourraient être des indicateurs précoces de la maladie, ouvrant ainsi la voie à des systèmes de détection basés sur la parole.

Dans ce contexte, l'objectif principal de cette thèse est de concevoir un système innovant pour évaluer les troubles du mouvement chez les patients atteints de la maladie de Parkinson à un stade précoce, en se basant sur l'analyse des signaux vocaux. Ce système devrait non seulement faciliter le diagnostic précoce de la maladie, mais également servir de support pour le suivi de la progression de la condition chez les patients déjà diagnostiqués.

Pour atteindre cet objectif, plusieurs sous-objectifs sont identifiés, notamment la construction d'une

base de données de patients atteints de la maladie de Parkinson pour l'entraînement et l'évaluation du système, la conception et l'implémentation d'algorithmes d'apprentissage automatique et d'apprentissage en profondeur adaptés à l'analyse des données de la maladie, ainsi que l'évaluation de la robustesse et de la généralisation des modèles de détection développés.

Le paragraphe précédent met en évidence l'importance d'avoir un système de détection précoce de la maladie de Parkinson (MP) efficace, économique, facilement accessible et non invasif. En effet, un tel système devrait améliorer les services de santé et assurer un suivi rapproché des patients. Un tel système peut utiliser la parole pour détecter précocement et surveiller la progression de la MP.

L'objectif principal de cette thèse est de construire un système pour évaluer les troubles du mouvement chez les patients atteints de la maladie de Parkinson à un stade précoce, en se basant sur les signaux vocaux, en utilisant des techniques d'apprentissage automatique de plus en plus populaires dans le domaine de la biologie, le système résultant servira de support pour le diagnostic différentiel des nouveaux patients atteints de la maladie de Parkinson.

Parmi Les sous-objectifs de cette thèse sont les suivants :

- Le système cible est basé sur les données. Nous nous appuyons sur des ensembles de données de patients atteints de la maladie de Parkinson pour construire, évaluer et comparer des solutions. À cette fin, nous avons utilisé une base de données de patients atteints de la maladie de Parkinson publiée par un hôpital, qui comprend des enregistrements de parole. Étant donné que notre objectif est la détection précoce de la maladie et que la collecte d'une grande base de données aux premiers stades de la maladie est très difficile, nous supposons que nous divisons la base de données des patients en deux groupes, où nous prenons un groupe pour l'étude et comparons les résultats obtenus avec les résultats initiaux pour évaluer l'efficacité du système.
- Concevoir et mettre en œuvre des algorithmes d'apprentissage automatique et d'apprentissage en profondeur spécifiquement adaptés à l'analyse des données de la maladie de Parkinson, en tenant compte des particularités des différents types de signaux et en optimisant les performances du système de détection.
- Évaluer la robustesse et la généralisation des modèles de détection développés en utilisant différentes métriques d'évaluation, telles que la sensibilité, la spécificité et la précision, ainsi que des techniques de validation croisée pour vérifier la validité des résultats obtenus sur des ensembles de données indépendants

Chapitre 1

La maladie de parkinson

1. Introduction

Ce chapitre présente la maladie de Parkinson, ses origines physiologiques, les symptômes induits, les échelles de notation utilisées pour évaluer le stade et la progression de la maladie, le diagnostic clinique, les médicaments possibles et leurs effets secondaires sur les mouvements, ainsi que la manière dont cette maladie peut affecter l'écriture, la parole et les mouvements oculaires même aux stades précoces. Il met également en lumière la difficulté de détecter la maladie aux stades précoces et l'importance de construire un système efficace et non invasif pour la détection précoce ; où la détection précoce peut être énormément bénéfique pour ralentir la progression de la maladie de Parkinson. Par conséquent, ce chapitre fournira au lecteur un aperçu de la condition et des raisons de rechercher de nouveaux schémas pour aider les cliniciens dans le diagnostic et l'évaluation.

2. La maladie Parkinson :

La maladie de Parkinson est aujourd'hui la deuxième maladie neurologique la plus courante causant un handicap après l'accident vasculaire cérébral et la maladie d'Alzheimer [Dubayová et al., 2010]. La maladie survient plus fréquemment chez les hommes que chez les femmes à chaque décennie de la vie, ce qui s'explique par les effets neuroprotecteurs des œstrogènes. La prévalence et l'incidence de la maladie de Parkinson dans les pays européens ont été estimées à environ 108 à 257 / 100 000 et 11 à 19 / 100 000 par an, respectivement. La prévalence dans les pays asiatiques est légèrement inférieure, la prévalence tous âges confondus variant de 51,3 à 176,9 / 100 000 personnes et l'incidence de 6,7 à 8,7 pour 100 000 personnes par an. Les taux de prévalence et d'incidence sont les plus bas dans les pays africains - la prévalence brute variant de 7 à 31,4 / 100 000 personnes et le taux d'incidence brute de la maladie de Parkinson étant de 4,5 / 100 000 personnes par an [Dubayová et al., 2010].

Les symptômes caractéristiques classiques de la maladie de Parkinson sont le tremblement au repos, la rigidité musculaire, l'akinésie (un retard dans le début des mouvements avec de longs temps de réaction), la bradykinésie (lenteur des mouvements) et la stabilité posturale, Changements dans la parole et l'écriture (La voix peut devenir plus douce, monotone ou tremblante, et l'écriture peut devenir plus petite et plus difficile à contrôler), En plus de ces symptômes moteurs, la maladie de Parkinson peut également entraîner des symptômes non moteurs tels que la dépression, l'anxiété, les problèmes de sommeil, les troubles cognitifs et la constipation . Le médicament le plus couramment prescrit pour traiter les patients atteints de la maladie de Parkinson est la lévodopa (L-dopa), puisque

la perte de neurones dopaminergiques dans le noyau substantiel du cerveau est associée à l'apparition des principaux symptômes moteurs de la maladie de Parkinson. L'étiologie de la maladie de Parkinson est également liée à l'âge et à l'exposition aux radicaux libres et aux toxines externes, ainsi qu'aux mutations génétiques [Weiner et al., 2013].

3. Pathophysiologie

La maladie de Parkinson se caractérise par une atteinte des neurones (cellules nerveuses) de la région du cerveau connue sous le nom de substance noire (du latin pour "substance noire") (Figure 2.1).

Lorsque les neurones de la substance noire dégénèrent, la capacité du cerveau à générer des mouvements corporels est perturbée, ce qui produit des signes et des symptômes caractéristiques de la maladie de Parkinson qui seront décrits dans la section 2.4 [Weiner et al., 2013]. Substance noire est une petite zone située en profondeur dans le cerveau. Il y a une substance noire du côté droit du cerveau et une du côté gauche, et souvent un côté est affecté avant l'autre. Pour cette raison, les personnes atteintes de la maladie de Parkinson présentent souvent des symptômes principalement d'un côté de leur corps, notamment aux stades précoces [Weiner et al., 2013]. Les symptômes de la maladie de Parkinson ne deviennent perceptibles que lorsqu'environ 80 % des cellules de la substance noire sont mortes, car le système nerveux humain comporte de multiples facteurs de sécurité et des redondances intégrées. Pendant longtemps, ces facteurs de sécurité ont été capables de prendre en charge les activités des cellules mourantes [Weiner et al., 2013].

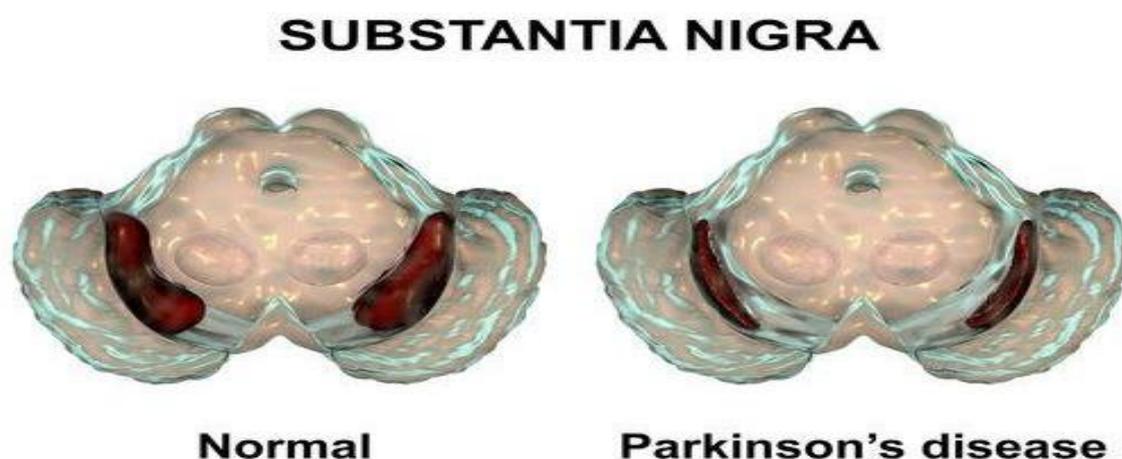


Figure 1 : Location of substance noire et la différence entre normal et la maladie parkinson .

Dans les échantillons de tissu cérébral prélevés sur des individus atteints de la maladie de Parkinson, le cerveau semble globalement normal, à l'exception de la perte de la pigmentation noire habituelle de la substantia nigra. La caractéristique distinctive de la maladie de Parkinson depuis des décennies a été la perte des cellules productrices de dopamine dans cette région du cerveau, accompagnée de la présence de grappes de protéines alpha-synucléine (connues sous le nom de corps de Lewy) [Weiner et al., 2013]. Bien que la substantia nigra représente une infime proportion du poids du cerveau, son rôle est crucial dans le contrôle de nos mouvements en raison de ses importantes connexions électrochimiques avec les centres moteurs du cerveau [Weiner et al., 2013]. Plus précisément, une série d'événements électriques et chimiques complexes à l'intérieur du cerveau transmet des informations d'un neurone à l'autre. Les substances chimiques que les cellules cérébrales utilisent pour communiquer entre elles sont appelées neurotransmetteurs ou, plus généralement, des neurochimiques. Le neurotransmetteur spécifique produit et utilisé par la substantia nigra est la dopamine [Weiner et al., 2013]. Lorsque les cellules de la substantia nigra se dégradent et meurent, la dopamine est perdue, et les messages relayés par la dopamine vers d'autres centres moteurs ne peuvent pas être transmis, ce qui entraîne un dysfonctionnement de la substance grise centrale [Weiner et al., 2013].

4. la progression de la maladie parkinson :

Dans la maladie de Parkinson, d'autres régions du cerveau au-delà de la substantia nigra sont impliquées dans l'évolution de la maladie. Des chercheurs ont constaté que des zones du tronc cérébral sous la substantia nigra présentent une perte de cellules et contiennent des agrégats de protéine alpha-synucléine, qui peuvent se former bien avant ceux de la substantia nigra [Port, 2019]. Par conséquent, Braak et ses collègues [Braak et Tredici, 2017] ont élaboré un système de stadification caractérisant la progression de la maladie de Parkinson. Ce système est divisé en six étapes distinctes, chacune étant associée à une pathologie anormale dans des structures neurologiques spécifiques. La pathologie liée à l'alpha-synucléine apparaît d'abord dans le bulbe olfactif et les régions du tronc cérébral, entraînant des symptômes non moteurs tels qu'une altération de l'odorat ou la constipation. Avec la progression de la maladie, la pathologie se propage au mésencéphale et au prosencéphale basal, causant des symptômes moteurs avant d'atteindre enfin le cortex, où les symptômes cognitifs se manifestent [Petersen, 2017]. Ce modèle suggère que la progression caractéristique des symptômes de la maladie de Parkinson correspond aux régions infiltrées par la pathologie de Lewy (Figure 2.2).

Dans la maladie de Parkinson, d'autres régions du cerveau au-delà de la substantia nigra sont impliquées dans l'évolution de la maladie. Des chercheurs ont constaté que des zones du tronc cérébral sous la substantia nigra présentent une perte de cellules et contiennent des agrégats de protéine alpha-synucléine, qui peuvent se former bien avant ceux de la substantia nigra [Port, 2019]. Par conséquent, Braak et ses collègues [Braak et Tredici, 2017] ont élaboré un système de stadification caractérisant la progression de la maladie de Parkinson. Ce système est divisé en six étapes distinctes, chacune étant associée à une pathologie anormale dans des structures neurologiques spécifiques. La pathologie liée à l'alpha-synucléine apparaît d'abord dans le bulbe olfactif et les régions du tronc cérébral, entraînant des symptômes non moteurs tels qu'une altération de l'odorat ou la constipation. Avec la progression de la maladie, la pathologie se propage au mésencéphale et au prosencéphale basal, causant des symptômes moteurs avant d'atteindre enfin le cortex, où les symptômes cognitifs se manifestent [Petersen, 2017]. Ce modèle suggère que la progression caractéristique des symptômes de la maladie de Parkinson correspond aux régions infiltrées par la pathologie de Lewy (Figure 2.2).

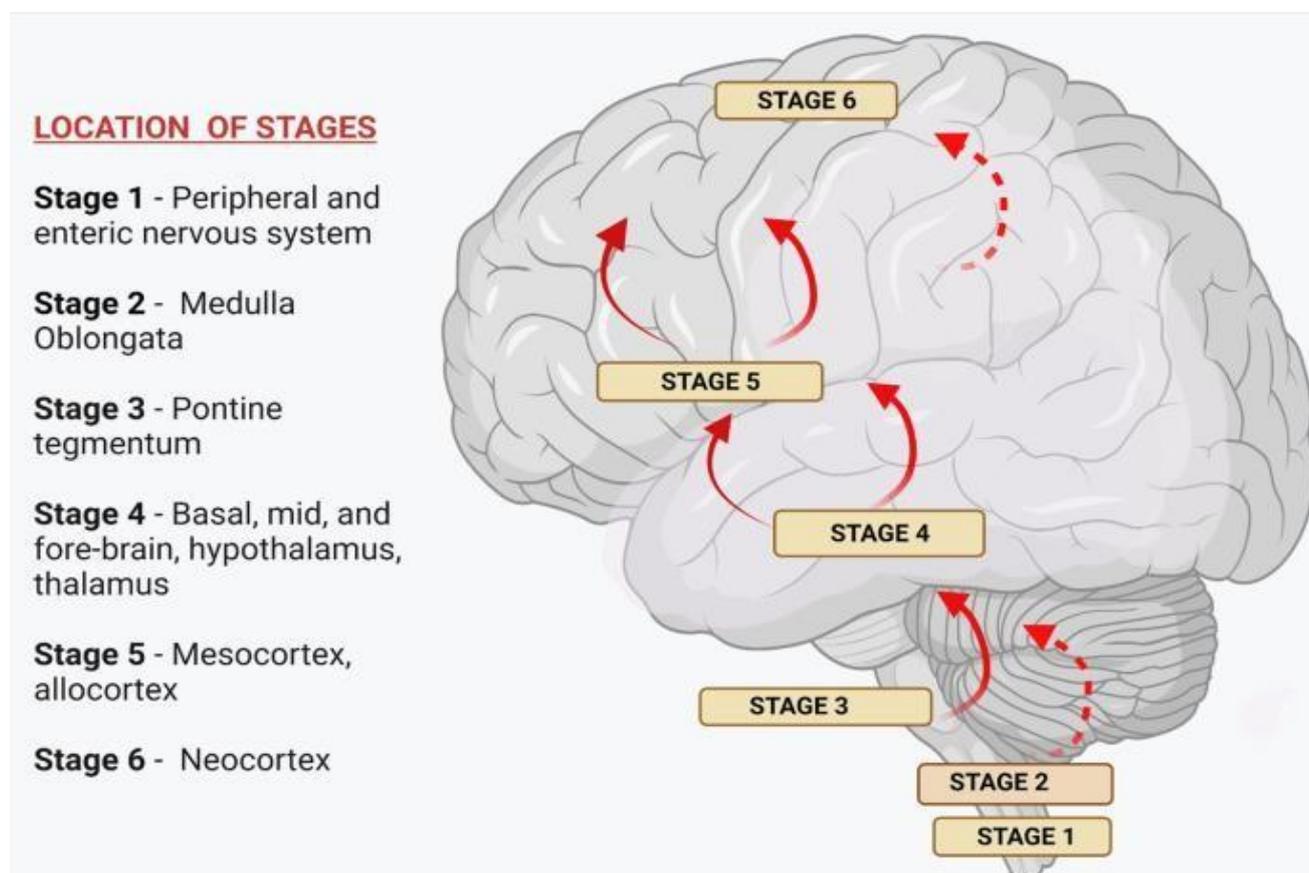


Figure 2 : Illustration de la stadification de Braak dans la maladie de Parkinson

5. La symptomatologie :

La maladie de Parkinson se manifeste par une grande variété de symptômes, dont beaucoup sont partagés avec d'autres conditions. C'est l'une des raisons pour lesquelles un diagnostic précis est compliqué et nécessite de longues périodes d'observation. Comme mentionné dans la section 2.2, la dopamine est tellement importante pour le contrôle du système nerveux central sur les muscles que lorsque ce neurotransmetteur est perdu, le tonus musculaire est altéré [Weiner et al., 2013].

Les symptômes caractéristiques de la maladie de Parkinson sont répartis en symptômes moteurs et non moteurs.

- Parmi les symptômes moteurs, les plus courants sont [Weiner et al., 2013]:

1. Akinesia : un retard dans le début des mouvements avec des temps de réaction prolongés.

2. Bradykinésie : lenteur des mouvements.

3. Hypokinésie : des mouvements pauvres, incomplets ou simplifiés. Dans la plupart des cas, les membres concernés n'atteignent pas l'amplitude complète associée à l'action ou au but du mouvement. Par exemple, l'écriture peut devenir plus petite.

4. Raideur : raideur des muscles, causant des douleurs et entravant certaines actions et postures.

5. Tremblement : il peut s'agir d'un tremblement postural ou de repos. De la même manière, il peut être constant ou intermittent. Dans certains cas, il s'agit d'un tremblement interne, non visible mais assez perturbant pour le patient.

6. Instabilité posturale : impliquant des problèmes d'équilibre en position debout ou en marchant.

7. Dystonie : torsions involontaires et répétitives et contractions musculaires soutenues.

- En ce qui concerne les symptômes non moteurs, les plus courants sont [Weiner et al., 2013]:

1. Dépression et anxiété. Cela peut impliquer l'incapacité du patient à exprimer ses émotions.

2. Constipation.

3. Perte d'odorat.

4. Problèmes de communication.

5. Démence et autres problèmes cognitifs tels que les hallucinations ou les difficultés de concentration et d'exécution de tâches complexes.

6. Problèmes de sommeil.

7. Perturbations sexuelles.

Dans la maladie de Parkinson, les premiers symptômes diffèrent légèrement de l'état normal et progressent lentement, peut-être sur plusieurs décennies. Ces symptômes peuvent se manifester à des degrés et des combinaisons variables selon les individus [Weiner et al., 2013].

a. Le diagnostic et les échelles de notation :

Un dépistage précoce de la maladie de Parkinson pourrait améliorer et maintenir la qualité de vie des patients et augmenter leur espérance de vie. Cependant, le fait est que diagnostiquer la maladie de Parkinson sur la base de symptômes très précoces est rarement possible. De plus, il n'existe pas de méthodes efficaces et fiables capables d'assurer un diagnostic précoce de la maladie de Parkinson avec certitude [Weiner et al., 2013]. Étant donné que les premiers symptômes peuvent être très subtils et peuvent être négligés ou confondus avec d'autres problèmes médicaux, un diagnostic ferme est difficile aux premiers stades. La seule façon de diagnostiquer avec précision est d'attendre et de voir si les symptômes typiques de la maladie de Parkinson s'aggravent [Weiner et al., 2013]. Un médecin qui pose un diagnostic de maladie de Parkinson doit avoir une expertise en diagnostic neurologique. Avec l'aggravation des symptômes de la maladie de Parkinson, un diagnostic définitif peut généralement être posé et des médicaments peuvent être prescrits pour aider l'individu à continuer à fonctionner [Weiner et al., 2013].

6. **Le diagnostic clinique :** lors de la réalisation d'un examen neurologique pour évaluer un patient présentant un trouble du mouvement, le médecin prend un historique et effectue un examen physique. Tout d'abord, le médecin interroge le patient et les membres de sa famille sur les symptômes, car les membres de la famille peuvent remarquer le tremblement initial de la maladie de Parkinson même avant que la personne affectée ne le fasse, et ils peuvent également repérer la posture et la démarche caractéristiques de la maladie de Parkinson. Ensuite, le médecin observe les mouvements des doigts, des mains et des pieds du patient, ainsi que sa marche, ses virages et sa résistance aux défis posturaux (en poussant le patient pour voir s'il peut éviter de tomber). Le médecin observe également le patient pour voir si des mouvements anormaux (par exemple, un tremblement) sont visibles et si le patient peut se lever facilement d'une chaise. L'expression faciale, les mouvements oculaires et la parole sont également examinés. Le médecin fléchit et étend ensuite le cou, les bras, les poignets et les jambes du patient pour rechercher un tonus musculaire anormal. La force et la coordination des bras sont évaluées [Weiner et al., 2013]. La fonction cognitive (mentale) du patient est également évaluée. L'examen de l'état mental mini-mental (MMSE) est

recommandé pour mesurer l'impairment cognitif [Kurlowicz et Wallace, 1999]. Il est également utilisé pour estimer la gravité et la progression de l'impairment cognitif et pour suivre l'évolution des changements cognitifs chez un individu au fil du temps. Le MMSE est un questionnaire de 30 points, comprenant l'enregistrement (répétition de consignes nommées), l'attention et le calcul, le rappel, le langage, la capacité à suivre des commandes simples et l'orientation. Un score de 24 ou plus (sur 30) indique une cognition normale. En dessous de cela, les scores peuvent indiquer un impairment cognitif sévère (≤ 9 points), modéré (10-18 points) ou léger (19-23 points).

En plus du MMSE, plusieurs échelles de notation différentes peuvent être utilisées pour évaluer le stade et la progression de la maladie de Parkinson chez un individu. Les deux plus utilisées sont l'échelle de Hoehn et Yahr (H&Y) et l'échelle de notation unifiée de la maladie de Parkinson (UPDRS) [Perlmutter, 2009].

- L'échelle de Hoehn et Yahr comprend plusieurs stades dont les valeurs peuvent varier de 1 à 5. La valeur 1 implique que le patient présente des incapacités fonctionnelles faibles ou inexistantes, tandis que la valeur 5 implique que le patient est totalement dépendant. Cette échelle a été modifiée pour inclure trois niveaux intermédiaires (0, 1,5 et 2,5) pour prendre en compte le cours intermédiaire de la maladie, comme suit :

1. **Stade 0** : Aucun signe de maladie.

2. **Stade 1** : Maladie unilatérale.

3. **Stade 1,5** : Maladie unilatérale avec implication axiale supplémentaire.

4. **Stade 2** : Maladie bilatérale, sans altération de l'équilibre.

5. **Stade 2,5** : Maladie bilatérale légère, avec récupération à l'épreuve du tirage.

6. **Stade 3** : Maladie bilatérale légère à modérée ; instabilité posturale partielle ; indépendant physiquement.

7. **Stade 4** : Handicap sévère ; capable de marcher ou de se tenir debout sans assistance.

8. **Stade 5** : Confiné à un fauteuil roulant ou au lit sauf s'il est aidé.

7. Traitement :

Lorsqu'un individu est diagnostiqué avec la maladie de Parkinson, il devient possible de gérer efficacement les symptômes de cette affection par le biais de choix de mode de vie sains, de médicaments, et parfois même de chirurgie. Les besoins en traitement évoluent au fil de la progression de la maladie. À chaque étape, il demeure crucial pour le patient de maintenir une activité physique, d'adopter une alimentation équilibrée et de surveiller son bien-être mental. Un

traitement précoce à base de médicaments peut aider la plupart des personnes atteintes de la maladie de Parkinson à conserver un mode de vie actif et à poursuivre leurs activités professionnelles. Les médicaments sont ajustés tout au long de l'évolution de la maladie afin de garantir un meilleur contrôle des symptômes et d'éviter les effets secondaires majeurs. L'adaptation des médicaments peut être complexe, ce qui souligne l'importance de consulter un spécialiste des troubles du mouvement.

7.1 Médicaments pour les symptômes moteurs :

Il existe plusieurs classes de médicaments disponibles pour le traitement efficace des symptômes moteurs tout au long de l'évolution de la maladie de Parkinson (MP).

- **Carbidopa-lévodopa** : Le traitement le plus efficace de la MP est la combinaison de médicaments de carbidopa-lévodopa, destinée à augmenter les niveaux de dopamine dans le cerveau, qui sont déficients chez les personnes atteintes de la MP. La lévodopa, qui est convertie en dopamine dans le cerveau, augmente les niveaux de dopamine dans le cerveau. La lévodopa réduit le tremblement, la raideur et les mouvements lents chez les personnes atteintes de MP idiopathique. La carbidopa empêche la dégradation de la lévodopa dans le corps avant qu'elle n'atteigne le cerveau. Par conséquent, l'ajout de carbidopa permet à la lévodopa d'atteindre plus efficacement le cerveau.

Les effets secondaires du traitement par carbidopa-lévodopa comprennent des nausées, une hypotension orthostatique, une somnolence (qui peut être soudaine, appelée attaques de sommeil), des comportements impulsifs, des hallucinations et une confusion. La lévodopa contribue également au développement de la dyskinésie (mouvements incontrôlés) [Golbe et al., 2012].

- **Infusion de carbidopa-lévodopa** : Une forme alternative de carbidopa-lévodopa a été approuvée par l'administration des aliments et des médicaments (FDA) en 2015. Elle est destinée aux personnes atteintes d'une maladie plus avancée, dont les symptômes ne répondent plus bien à la carbidopa-lévodopa par voie orale. Au lieu de prendre un comprimé, les personnes atteintes de la MP peuvent recevoir de la carbidopa-lévodopa sous forme de gel par le biais d'une pompe à perfusion. La pompe délivre le médicament directement dans l'intestin grêle par le biais d'un tube placé chirurgicalement. L'avantage d'une perfusion continue de carbidopa-lévodopa est une réduction de l'immobilité ou du temps "off" (lorsque les médicaments ne suffisent pas à maintenir un bon contrôle des symptômes tout au long de la journée) de la lévodopa.

Les effets secondaires de la perfusion de carbidopa-lévodopa sont similaires à ceux de la carbidopa-lévodopa par voie orale, mais peuvent être associés à une incidence plus élevée de neuropathie périphérique (engourdissement ou perte de sensation dans les doigts ou les pieds) [Golbe et al., 2012].

- **Agonistes de la dopamine** : sont un peu différents de la carbidopa-lévodopa en ce sens qu'au lieu d'augmenter les niveaux de dopamine dans le cerveau, ils imitent l'activité de la dopamine. Ils peuvent être administrés seuls dans les premiers stades de la MP, ou en complément de la carbidopa-lévodopa ou d'autres médicaments contre la MP.

Les effets secondaires des agonistes de la dopamine sont similaires à ceux de la carbidopa-lévodopa, bien que les troubles du contrôle des impulsions et l'apparition soudaine de somnolence puissent être plus prononcés [Golbe et al., 2012].

- **Inhibiteurs de la catéchol-O-méthyltransférase (COMT)** : sont parfois utilisés avec la carbidopa-lévodopa. Comme la carbidopa, ils empêchent la dégradation de la lévodopa avant qu'elle n'atteigne le cerveau. Le résultat est qu'un approvisionnement plus fiable en lévodopa pénètre dans le cerveau, où elle peut être convertie en dopamine. Les inhibiteurs de la COMT sont généralement prescrits pour traiter les "off" fréquents avec la thérapie à la lévodopa. Parfois, les inhibiteurs de la COMT peuvent augmenter les effets secondaires associés à la thérapie à la lévodopa.

Les autres effets secondaires courants des inhibiteurs de la COMT sont les douleurs abdominales, la diarrhée et la décoloration des liquides corporels tels que l'urine [Golbe et al., 2012].

- **Inhibiteurs sélectifs de la monoamine oxydase B (MAO-B)** : bloquent l'enzyme MAO-B dans le cerveau, qui dégrade la dopamine. Il s'agit d'un autre moyen d'augmenter les niveaux de dopamine dans le cerveau. Les inhibiteurs de la MAO-B peuvent être utilisés seuls ou avec d'autres médicaments contre la MP. Les inhibiteurs sélectifs de la MAO-B peuvent être prescrits en complément de la thérapie à la carbidopa-lévodopa, en particulier si les personnes présentent des symptômes de "wearing-off" lors de la prise de lévodopa.

Les effets secondaires des inhibiteurs sélectifs de la MAO-B comprennent des nausées légères, une bouche sèche, des étourdissements, une constipation et, parfois, des hallucinations et une confusion [Golbe et al., 2012].

- **Les anticholinergiques** : sont souvent utilisés pour la gestion de la maladie de Parkinson en tant que médicaments adjuvants à d'autres thérapies pour la maladie de Parkinson. Les

anticholinergiques sont souvent prescrits pour réduire le tremblement caractéristique de la maladie de Parkinson ou pour soulager les problèmes associés à la diminution de l'effet de la lévodopa.

Les effets secondaires courants des anticholinergiques incluent la confusion, les hallucinations, la constipation, la bouche sèche et les problèmes urinaires. Par conséquent, l'utilisation d'anticholinergiques est généralement limitée aux personnes plus jeunes atteintes de la maladie de Parkinson (âgées de moins de 70 ans). Ces anticholinergiques doivent également être évités en combinaison avec des antihistaminiques, certains médicaments psychiatriques et de l'alcool [Golbe et al., 2012].

- **Les formulations d'amantadine** : ont été observées pour soulager le tremblement de la maladie de Parkinson ainsi que la rigidité musculaire. Elle est généralement utilisée comme médicament adjuvant à d'autres thérapies pour la maladie de Parkinson. De plus, elle est utilisée pour réduire la dyskinésie ou les mouvements involontaires causés par la lévodopa et pour réduire le "temps d'arrêt".

Les effets secondaires courants incluent des vertiges, une bouche sèche, une constipation, des rêves vifs, une éruption cutanée en dentelle, généralement sur les jambes, et un gonflement des chevilles. Elle peut également interagir avec ou renforcer les effets secondaires des anticholinergiques et de la thérapie par lévodopa [Golbe et al., 2012].

- **Les inhibiteurs de l'adénosine** : bloquent les effets du récepteur de l'adénosine. Comme la dopamine, l'adénosine est un neurotransmetteur qui agit dans le striatum, les structures profondes du cerveau affectées dans la maladie de Parkinson. Cependant, dans une certaine mesure, l'adénosine et la dopamine ont des effets opposés, de sorte que l'inhibition du récepteur de l'adénosine améliore la fonction motrice. Il est indiqué comme traitement complémentaire à la lévodopa pour ceux qui éprouvent des épisodes de "hors période" [Golbe et al., 2012].

7.2 Traitement chirurgical :

- **Deep Brain Stimulation /La stimulation cérébrale profonde (SCP)** est une procédure neurochirurgicale destinée aux personnes atteintes de la maladie de Parkinson avancée qui conservent une bonne réponse à la lévodopa, mais qui ont développé des fluctuations motrices significatives, notamment la dyskinésie. La SCP peut également être utilisée pour traiter le tremblement résistant aux médicaments. En stimulant des points spécifiques dans les circuits de contrôle moteur du cerveau, la SCP "rééquilibre" les circuits, restaurant dans une

certaine mesure le contrôle normal des mouvements. Dans la plupart des cas, cela permet à la personne atteinte de la maladie de Parkinson de réduire sa dose de lévodopa et donc de réduire sa dyskinésie tout en maintenant un bon contrôle des symptômes. La SCP implique l'implantation d'électrodes permanentes et fines dans des parties profondes sélectionnées du cerveau. Un générateur d'impulsions alimenté par batterie, similaire à un stimulateur cardiaque, est implanté sous la peau de la poitrine ou de l'abdomen. Le générateur d'impulsions est connecté aux électrodes de stimulation via des fils, qui sont tunnelisés sous la peau du cuir chevelu et du cou. Les électrodes de stimulation sont programmées par un ordinateur distant pour un contrôle maximal des symptômes, et les batteries peuvent être remplacées par une procédure ambulatoire lorsque nécessaire, généralement après plusieurs années.

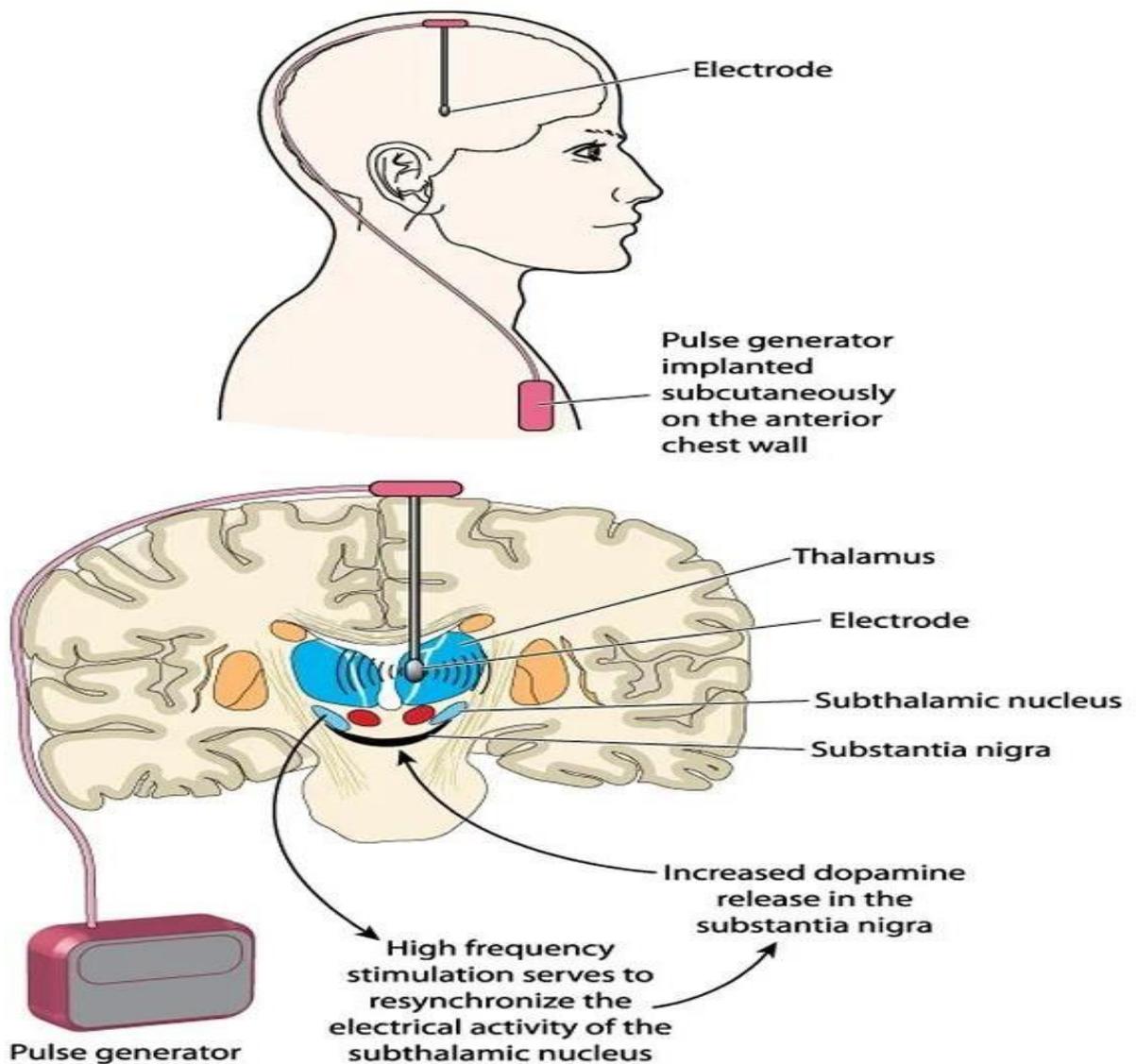


Figure 3 : Deep Brain Stimulation

- Stem cell transplantations /Greffe de cellules souches** : Les traitements actuellement disponibles pour traiter la maladie de Parkinson se concentrent sur l'augmentation des niveaux de dopamine dans le cerveau, ce qui soulage les symptômes moteurs. Cependant, l'efficacité de ces traitements diminue avec le temps et est associée à divers effets secondaires [Henchcliffe et Parmar, 2018]. Étant donné que les médicaments ne vont qu'en partie pour traiter pleinement les problèmes de coordination et de mouvement, un traitement possible à long terme pour la maladie de Parkinson consiste à greffer des cellules souches produisant de la dopamine dans le cerveau des patients. Cette approche pourrait entraîner un soulagement à long terme des symptômes moteurs et peut réduire ou même arrêter le besoin de médicaments. L'utilisation de cellules souches fabriquées en laboratoire pour les greffes permettrait de produire suffisamment de cellules pour répondre à la demande actuelle, de garantir que le nombre et le type de cellules greffées sont toujours les mêmes, qu'elles produisent les quantités de dopamine souhaitées, et que le risque de rejet est faible. Les cellules utilisées pour effectuer les greffes sont extraites du sang ou de la peau du patient, puis traitées pour devenir presque n'importe quel type de cellule dans le corps [Henchcliffe et Parmar, 2018].

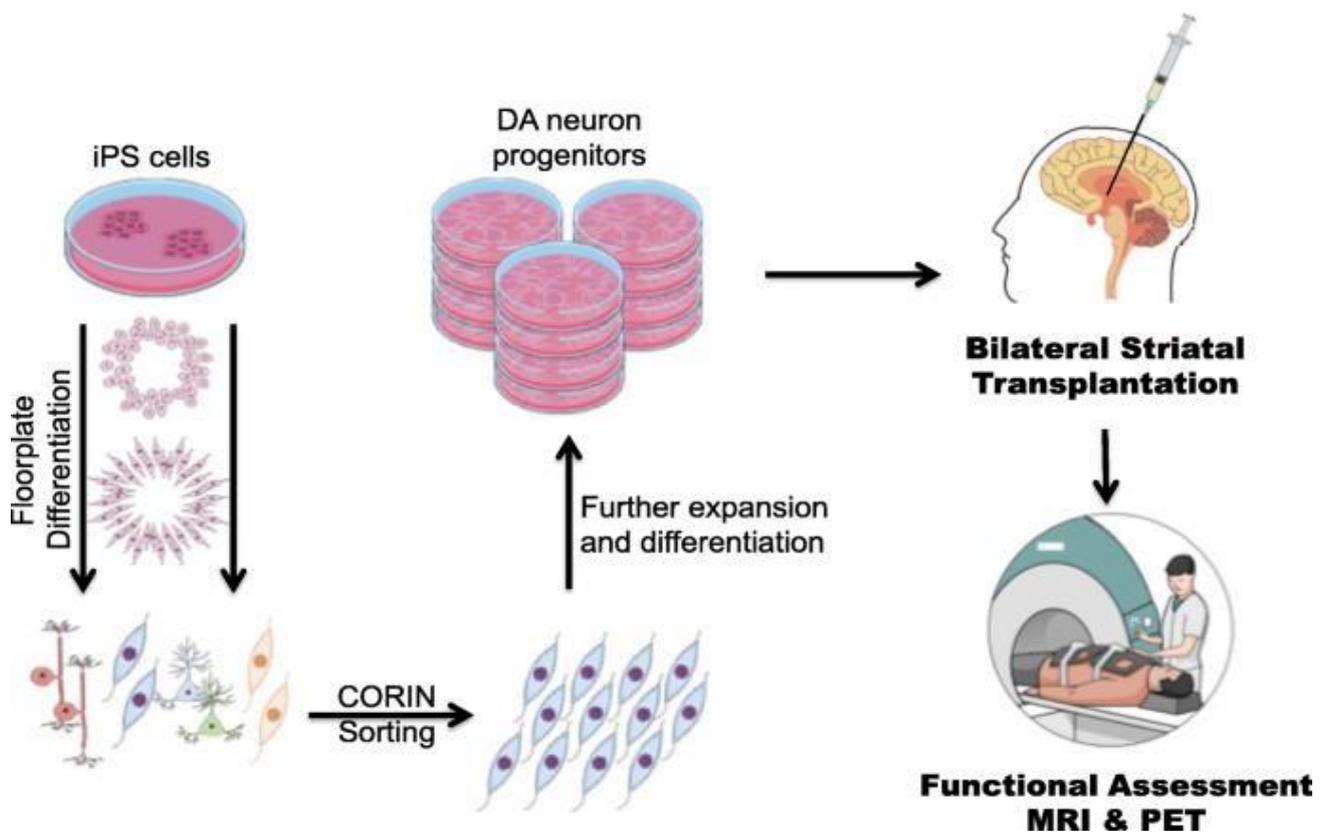


Figure 4 : Stem cell transplantations

8. La maladie de Parkinson et la parole :

La parole requiert la coordination harmonieuse de plusieurs activités neurologiques et musculaires, incluant la cognition, le mouvement, et la coordination musculaire (Davie, 2013). Ces processus, regroupant la planification, la programmation, le contrôle, et l'exécution de la parole, sont collectivement désignés sous le terme de "processus moteurs de la parole".

Lorsque le système nerveux est perturbé, la production de la parole peut en souffrir, signalant parfois une maladie neurologique. Les troubles de la parole liés à la maladie de Parkinson, connus sous le nom de dysarthrie, se manifestent par une altération de la clarté de la parole. La dysarthrie reflète des altérations neuro-musculaires affectant la force, la vitesse, le ton, ainsi que la stabilité ou la précision des mouvements impliqués dans la parole.

Ces altérations touchent également les composantes essentielles de la mécanique de la parole, incluant la respiration, la voix, les vibrations, l'articulation, et l'intonation (Davie, 2013).

Concernant la maladie de Parkinson et ses répercussions sur la parole, il est à noter un impact significatif sur l'articulation et la communication, dû en partie aux modifications induites par la maladie au niveau du système nerveux et des muscles.

Parmi les effets fréquemment observés de la maladie de Parkinson sur la parole, on compte les difficultés d'élocution, pouvant rendre la parole peu claire ou incompréhensible en raison d'un contrôle musculaire insuffisant des articulations nécessaires à la prononciation. De nombreuses personnes atteintes de la maladie de Parkinson présentent également une faiblesse vocale, caractérisée par une voix faible et peu audible.

Des altérations du ton et du rythme sont également observées, impactant ainsi l'expression et l'interaction avec autrui. La modulation tonale peut devenir moins dynamique et le rythme plus lent et monotone, donnant parfois l'impression d'une parole hachée ou dénuée de vivacité.

Ces changements peuvent considérablement affecter la qualité de vie quotidienne et la capacité à communiquer efficacement avec les autres. Toutefois, diverses stratégies et traitements sont disponibles pour aider les patients à faire face à ces défis, notamment la thérapie logopédique visant à améliorer la prononciation et la communication.

9. Conclusion :

En résumé, la maladie de Parkinson est une condition neurologique chronique qui affecte le mouvement, l'équilibre et la coordination musculaire. Cette maladie entraîne progressivement une diminution de la production de dopamine dans le cerveau, ce qui se traduit par des symptômes tels que les tremblements, la rigidité musculaire et des difficultés de coordination motrice. La progression de la maladie peut entraîner une détérioration de l'équilibre et une augmentation de la faiblesse musculaire, ce qui nuit à la qualité de vie. Le traitement de la maladie de Parkinson comprend des médicaments, de la physiothérapie, de la thérapie comportementale et, dans certains cas graves, une intervention chirurgicale. Les recherches actuelles se concentrent sur le développement de nouveaux traitements pour améliorer la qualité de vie des personnes atteintes de la maladie de Parkinson.

Il est clair que la maladie de Parkinson a un impact significatif sur la capacité de parler et de communiquer. Les troubles articulatoires, la faiblesse vocale et les changements de tonalité et de rythme sont des manifestations courantes de cette maladie. Ces altérations peuvent entraîner une diminution de l'intelligibilité de la parole et affecter la qualité de vie quotidienne ainsi que la capacité à interagir avec les autres.

Il est important de reconnaître que la parole est un processus complexe qui implique la coordination de nombreuses activités neurologiques et musculaires. Lorsque le système nerveux est perturbé, comme c'est le cas dans la maladie de Parkinson, cela peut entraîner des dysfonctionnements dans les processus de planification, de programmation, de contrôle et d'exécution de la parole.

Cependant, malgré les défis posés par la maladie de Parkinson, il existe des stratégies et des traitements disponibles pour aider à améliorer la communication. La rééducation logopédique, notamment, peut jouer un rôle crucial dans l'amélioration de la qualité de la parole et dans le maintien d'une communication efficace.

En fin de compte, il est essentiel de sensibiliser aux difficultés rencontrées par les personnes atteintes de la maladie de Parkinson en matière de parole et de communication, et de continuer à rechercher des solutions innovantes pour répondre à leurs besoins.

Chapitre02

Apprentissage Automatique

1. Introduction :

- **Vue générale sur l'intelligence artificielle :**

L'intelligence artificielle (IA) peut être définie comme "la science et l'ingénierie de la création de machines intelligentes, en particulier de programmes informatiques intelligents", selon John McCarthy, considéré comme le père de l'IA.

Elle se matérialise par des programmes informatiques capables d'effectuer des opérations semblables à celles du cerveau humain. Le Deep Learning, une forme complexe d'apprentissage automatique, en est un exemple. Il s'agit d'un ensemble de méthodes qui tentent de transformer de manière non linéaire les données à travers différentes architectures articulées et niveaux d'abstraction.

L'apprentissage automatique et le Deep Learning font tous deux partie de l'intelligence artificielle, permettant aux ordinateurs de prendre des décisions judicieuses. On résume la relation entre les trois concepts d'IA, ML et DL comme suit :

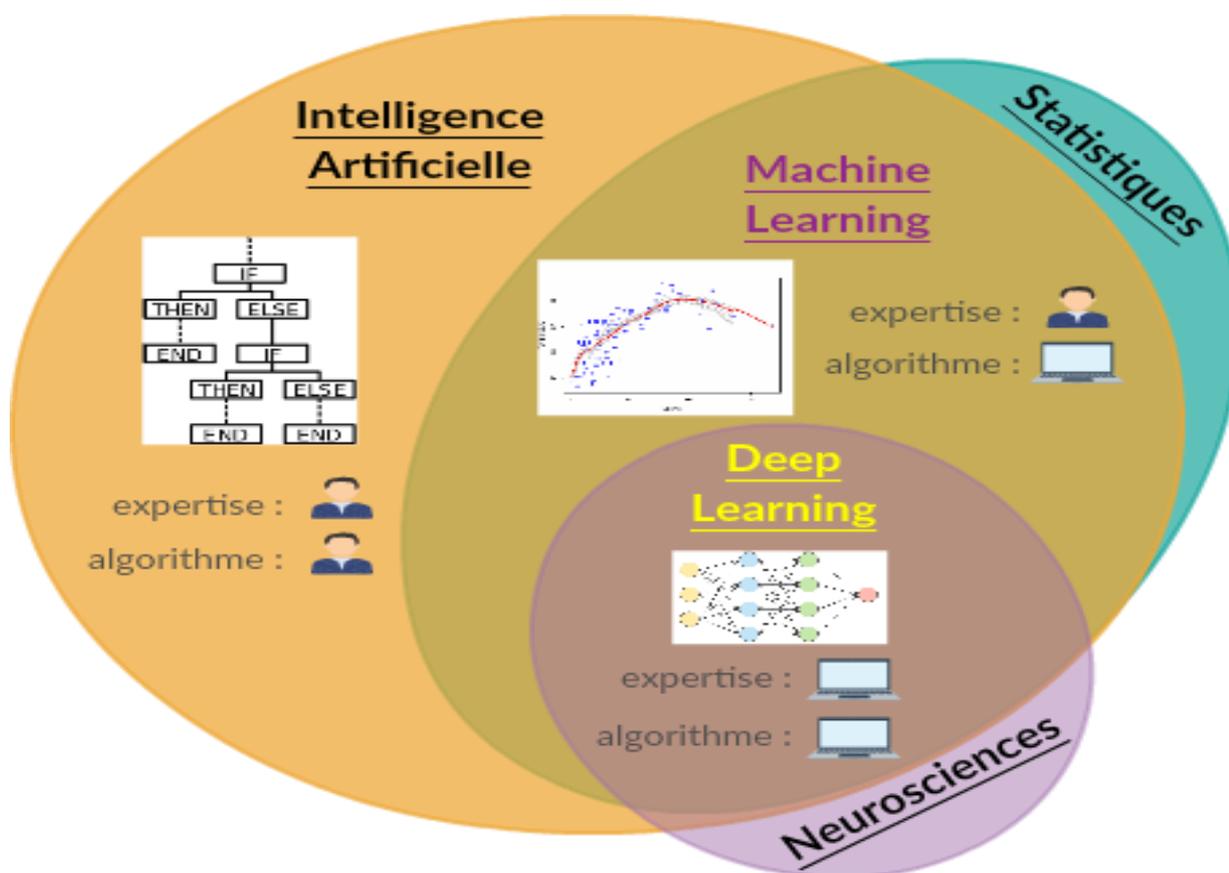


Figure 5 : La relation entre IA, ML et DL

L'idée en créant l'intelligence artificielle est de rendre la vie plus facile pour les humains.

L'intelligence artificielle est une conception de machine capable de penser. C'est l'intelligence de machine. Les chercheurs en intelligence artificielle veulent apporter des émotions aux machines en même temps que l'intelligence générale.

➤ **L'Objectif de l'intelligence artificielle est de :**

Création de systèmes experts : Les systèmes doivent être dotés de comportements intelligents, de capacités d'apprentissage, de démonstration, d'explication, et être en mesure de fournir aux utilisateurs les meilleurs conseils. Certaines applications intègrent la machine, la programmation et des données spécifiques pour faciliter la réflexion, offrant ainsi des éclaircissements et des incitations aux utilisateurs.

Implémentation de l'intelligence dans les machines : L'objectif est de développer des systèmes capables de comprendre, de penser, d'apprendre et de se comporter comme des humains.

Les robots intelligents et les applications médicales sont fortement influencés par l'intelligence artificielle. Un exemple populaire d'application médicale de l'IA est la radiochirurgie, utilisée dans les opérations tumorales pour minimiser les dommages aux tissus environnants. Elle est également utilisée dans la détection et la surveillance des troubles neurologiques en simulant les fonctions cérébrales.

Grâce à l'intelligence artificielle des machines, les médecins peuvent évaluer les patients et leurs risques pour la santé, ainsi que les informer des effets secondaires de divers médicaments.

Les robots sont parmi les meilleures inventions humaines, capables d'accomplir de multiples tâches rapidement. Bien qu'ils ne puissent pas remplacer les humains, ils sont très efficaces pour les tâches répétitives et simples. La robotique est souvent utilisée pour aider les patients atteints de troubles mentaux à surmonter la dépression et à maintenir leur activité.

2. Apprentissage automatique :

L'apprentissage automatique est une approche analytique des données qui se concentre sur le développement et l'évaluation d'algorithmes. Il s'agit d'une discipline qui permet aux ordinateurs d'agir sans programmation explicite. Cette méthode se caractérise par sa capacité à sélectionner des caractéristiques pertinentes pour la détection de modèles, la classification et la prédiction, en se basant sur les modèles dérivés des données disponibles.

3. L'impact de l'apprentissage automatique sur le domaine de la santé :

Accepteriez-vous de confier votre traitement à un ordinateur ou à un robot à l'hôpital, ou même de remplacer votre médecin généraliste par une machine ?

Il est probable que vous hésiteriez à répondre affirmativement. Pourtant, le secteur de la santé est l'un des nombreux domaines que l'intelligence artificielle s'apprête à révolutionner.

Le diagnostic revêt une importance capitale dans le domaine médical. Pour assurer un traitement efficace, il est crucial d'identifier avec précision la maladie du patient. Actuellement, les soins de santé personnalisés sont souvent réservés à une élite aisée. Cependant, grâce à l'analyse prédictive et à l'apprentissage automatique, ces soins pourraient devenir plus accessibles. En effet, l'apprentissage permet aux médecins de sélectionner le bon diagnostic parmi un éventail limité de possibilités, en se basant sur les informations génétiques disponibles sur le patient. Ainsi, les praticiens peuvent tirer des enseignements de l'historique médical complet du patient. De plus, les dispositifs connectés comme l'Apple Watch et ses mesures d'électroencéphalogramme fournissent une quantité accrue de données qui pourraient faciliter l'élaboration de traitements personnalisés pour chaque patient.

4. Les différents types d'apprentissage :

Il y a un large éventail d'algorithmes utilisés dans l'apprentissage automatique. Ils peuvent être regroupés en plusieurs catégories différentes.

➤ **Apprentissage Supervisé :**

L'apprentissage supervisé est incontestablement la méthode d'apprentissage automatique la plus répandue. Il s'agit de former le programme sur des exemples dont la catégorie est connue (d'où le terme "supervisé"). Le programme dispose d'une base de données contenant des classes connues. Son

objectif est d'associer la valeur qui lui est présentée à la classe préalablement définie. L'algorithme classera ensuite cette nouvelle valeur selon les catégories existantes. Bien qu'il ne soit pas conçu pour créer de nouvelles catégories, cet algorithme s'avère très efficace pour des tâches de classification spécifiques.

Parmi les algorithmes utilisés en apprentissage supervisé, on trouve les réseaux de neurones, les classifieurs bayésiens naïfs, la régression linéaire, la régression logistique, les arbres de décision, la méthode des k plus proches voisins (KNN), etc.

Il existe de nombreuses applications pour mettre en œuvre un apprentissage supervisé :

- la bio-informatique
- La reconnaissance vocale, des formes ou de l'écriture manuscrite
- Le traitement automatique de la langue
- La vision par ordinateur

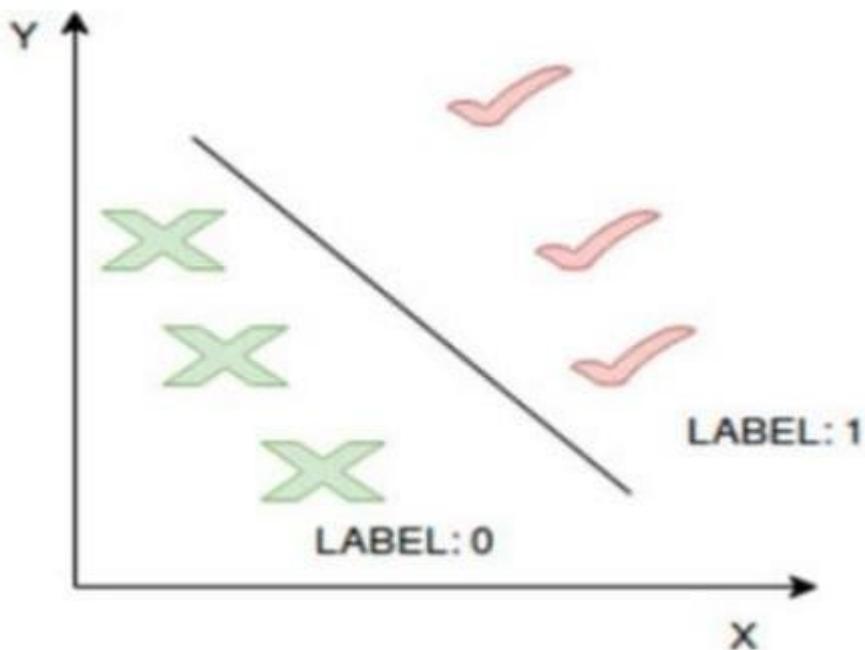


Figure 6: L'apprentissage supervisé

➤ **Apprentissage non-supervisé :**

Dans ce contexte, le programme est confronté uniquement à des valeurs et doit établir les catégories dans lesquelles les classer. Ainsi, il va "décider" lui-même du nombre de catégories à créer afin d'organiser les données en conséquence. Ces algorithmes sont employés lorsque nous ne disposons pas d'un échantillon préexistant.

On peut recourir à cette méthode pour constituer un échantillon, puis utiliser l'apprentissage supervisé pour classifier de nouvelles données. Ce type d'apprentissage est utile pour résoudre des problèmes de regroupement, de réduction de dimensionnalité ou d'apprentissage des règles d'association. Par exemple, les algorithmes Apriori et K-moyennes sont souvent utilisés à cet effet.

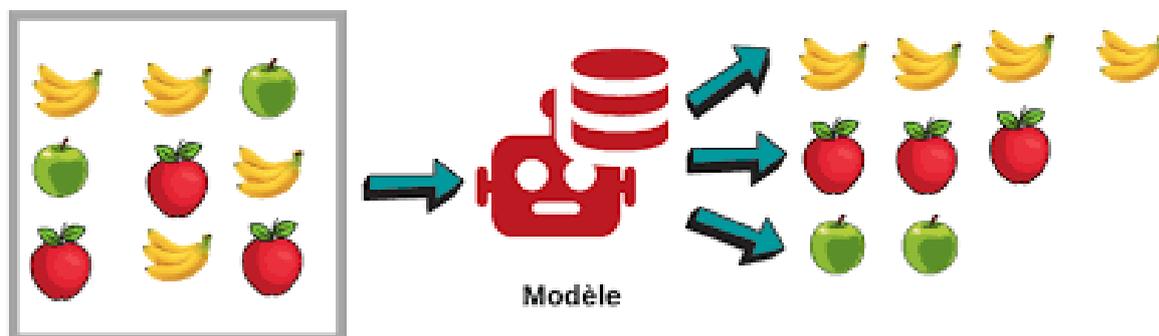


Figure 7 :L'apprentissage non-supervisé

➤ **la différence entre apprentissage supervisé et non supervisé :**

L'apprentissage supervisé utilise des données étiquetées pour entraîner un modèle à prédire des résultats, tandis que l'apprentissage non supervisé cherche à découvrir des structures intrinsèques dans des données non étiquetées.

L'apprentissage supervisé nécessite des données d'entraînement étiquetées, où chaque exemple est associé à une étiquette ou une réponse attendue. Le modèle est formé en utilisant ces exemples pour apprendre à prédire correctement les étiquettes pour de nouvelles données non étiquetées. En revanche, l'apprentissage non supervisé ne nécessite pas d'étiquettes dans les données d'entraînement. Il cherche

plutôt à trouver des schémas ou des structures intrinsèques dans les données sans référence à des étiquettes prédéfinies, en regroupant les données similaires ou en réduisant leur dimensionnalité.

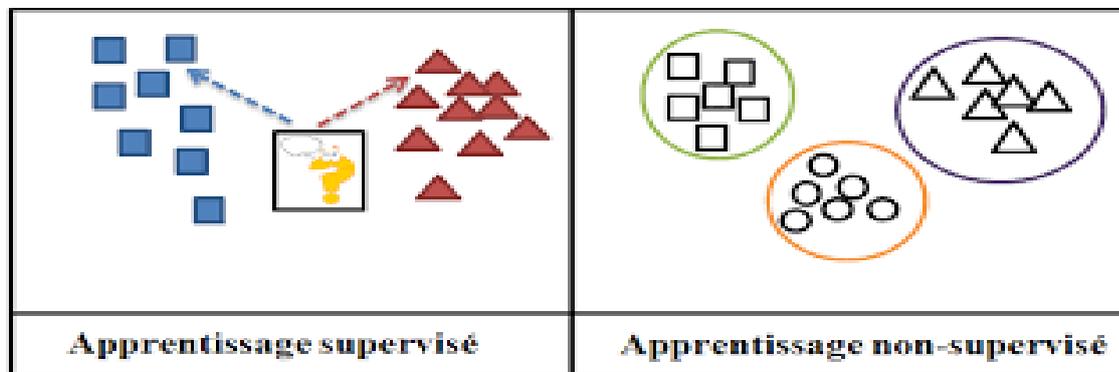


Figure 8 : Apprentissage supervisé vs non supervisé

➤ **Apprentissage semi-supervisé :**

L'apprentissage semi-supervisé est une approche de l'apprentissage automatique qui utilise à la fois des données étiquetées et non étiquetées pour l'entraînement des modèles. Contrairement à l'apprentissage supervisé traditionnel, où toutes les données sont étiquetées, et à l'apprentissage non supervisé, où aucune étiquette n'est fournie, l'apprentissage semi-supervisé combine les deux types de données.

Les méthodes d'apprentissage semi-supervisé incluent des techniques telles que la propagation de l'étiquette, où les étiquettes des données étiquetées sont propagées aux données non étiquetées en fonction de leur similarité, et les modèles générateurs-discriminateurs, où un générateur tente de créer des données réalistes pour tromper un discriminateur qui essaie de distinguer entre les données réelles et générées.

L'apprentissage semi-supervisé peut être appliqué à divers domaines, y compris la vision par ordinateur, le traitement du langage naturel, la reconnaissance de formes, et d'autres domaines où les données non étiquetées sont abondantes. En utilisant efficacement ces données non étiquetées, l'apprentissage semi-supervisé peut souvent améliorer les performances des modèles par rapport à l'apprentissage supervisé seul.

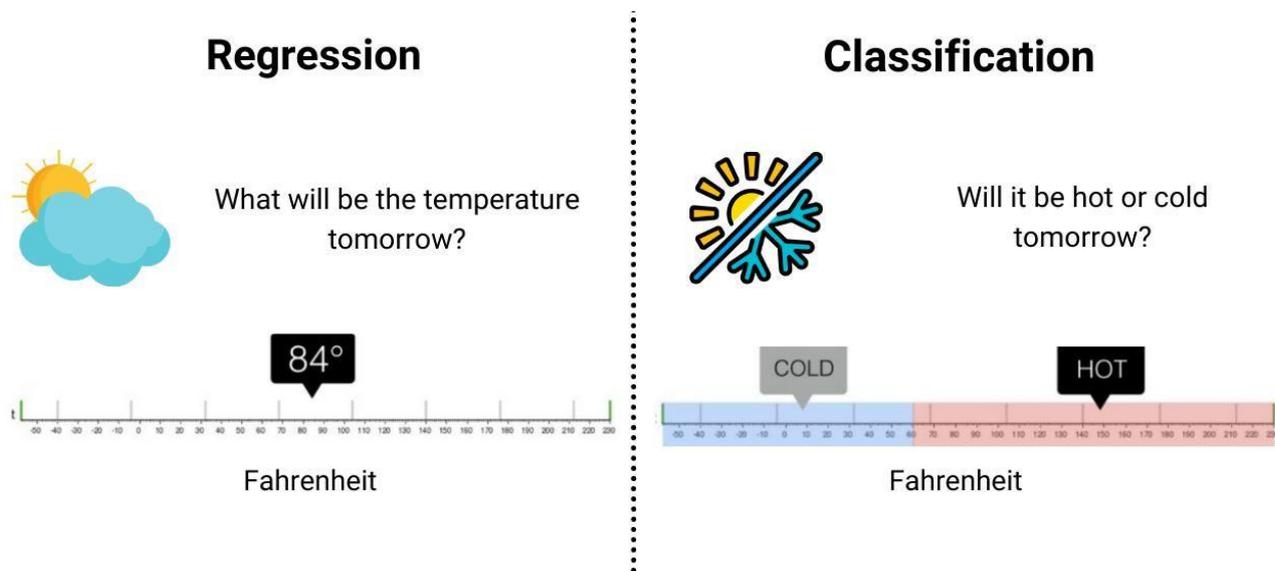


Figure. 9 : Classification vs. Régression

La classification et la régression sont deux types de modèles utilisés en apprentissage automatique pour résoudre différents types de problèmes.

- **Classification :**

Dans la classification, le modèle prédit une catégorie ou une classe pour une observation donnée.

Les données de sortie sont discrètes et qualitatives, Dans la classification, le modèle cherche à séparer les données en différentes classes ou catégories en fonction des caractéristiques d'entrée.

- **Régression :**

En revanche, la régression est utilisée lorsque la variable de sortie est continue et quantitative. Les données de sortie sont des valeurs numériques, exemples typiques de problèmes de régression comprennent la prédiction des prix des maisons en fonction de leurs caractéristiques, la prévision des ventes en fonction des dépenses publicitaires, ou l'estimation de la consommation d'énergie en fonction de divers facteurs.

En résumé, la classification est utilisée pour prédire des catégories ou des classes, tandis que la régression est utilisée pour prédire des valeurs numériques continues.

➤ **Apprentissage par renforcement :**

L'apprentissage par renforcement est une méthode d'apprentissage automatique où un agent apprend à prendre des décisions en interagissant avec un environnement. L'objectif de l'agent est de maximiser une récompense cumulée au fil du temps en choisissant les actions appropriées dans différentes situations.

L'apprentissage par renforcement est largement utilisé dans des domaines tels que les jeux, la robotique, la finance et les systèmes de recommandation, où les agents doivent prendre des décisions séquentielles pour atteindre des objectifs spécifiques.

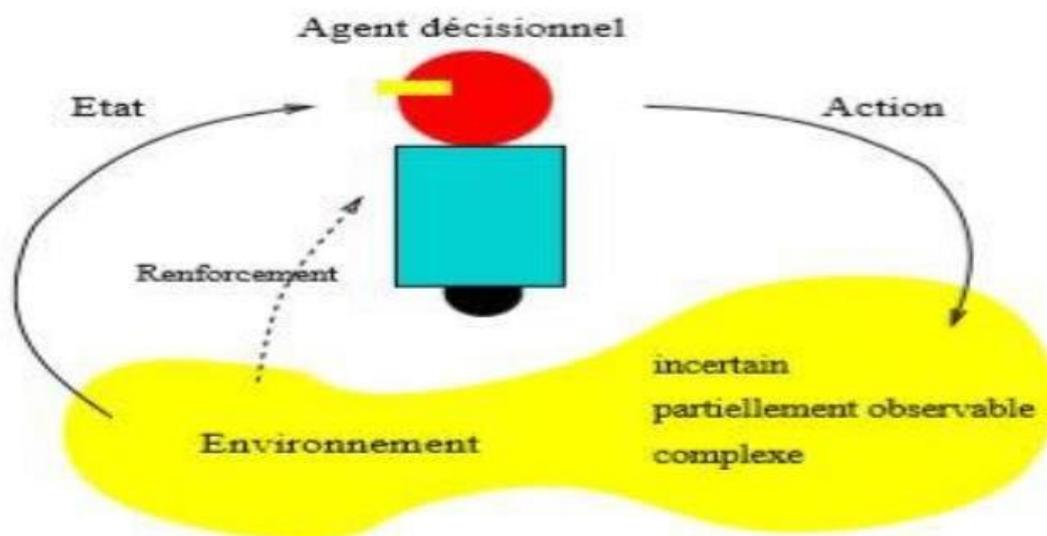


Figure 10 : L'apprentissage par renforcement

Ces familles d'algorithmes comptent parmi les plus utilisées pour l'apprentissage automatique. Il en existe cependant beaucoup d'autres, et certains algorithmes appartiennent à plusieurs catégories à la fois.

5. Algorithmes d'apprentissage automatique :

Les algorithmes d'apprentissage automatique sont des techniques informatiques qui permettent aux systèmes de "lire" les données et d'en tirer des conclusions sans être explicitement programmés pour effectuer une tâche spécifique. Au lieu de cela, ils apprennent à partir des exemples et des modèles présents dans les données.

Ces algorithmes sont devenus essentiels dans de nombreux domaines, notamment sur Internet. Par exemple, sur les réseaux sociaux comme Facebook, les algorithmes d'apprentissage automatique déterminent le contenu qui apparaît sur votre fil d'actualité en fonction de vos interactions passées et de celles de vos amis. De même, sur les services de messagerie comme Gmail, ces algorithmes filtrent les e-mails indésirables et classent les messages importants en se basant sur votre comportement de lecture.

En résumé, les algorithmes d'apprentissage automatique permettent aux systèmes informatiques de s'adapter et de fournir des recommandations personnalisées en se basant sur les données disponibles, ce qui améliore l'expérience utilisateur dans de nombreux domaines.

➤ **Algorithme SVM :**

Les SVM, ou "Support Vector Machines" en anglais, sont un algorithme d'apprentissage automatique supervisé largement utilisé dans les problèmes de classification, introduit par Vapnik en 1990. Cet algorithme consiste à représenter chaque donnée comme un point dans un espace de caractéristiques de dimension n et à les classer à l'aide d'un hyperplan qui sépare les classes.

Dans le SVM, chaque point de données est attribué à une classe spécifique, et l'objectif est de déterminer quelle classe convient le mieux à de nouvelles données. L'objectif principal du SVM est de trouver une frontière, qu'elle soit linéaire ou non linéaire, qui maximise la marge entre les différentes classes d'observations.

Bien que les modèles linéairement séparables soient simples à traiter et permettent de trouver rapidement un classificateur linéaire, dans la plupart des problèmes réels, il n'existe pas de frontière de séparation linéaire. C'est là que les SVM non linéaires entrent en jeu, qui sont capables de trouver des frontières de décision plus complexes pour des données non linéairement séparables.

En résumé, les SVM sont des outils puissants pour la classification, capables de gérer à la fois des problèmes de séparation linéaire et non linéaire, ce qui en fait un choix populaire dans de nombreux domaines de l'apprentissage automatique.

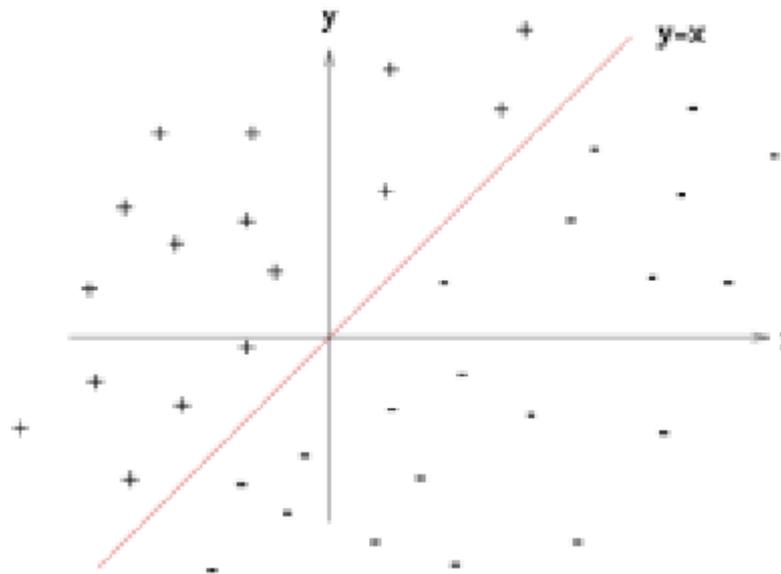


Figure 11 : classification à deux classes avec une séparatrice linéaire

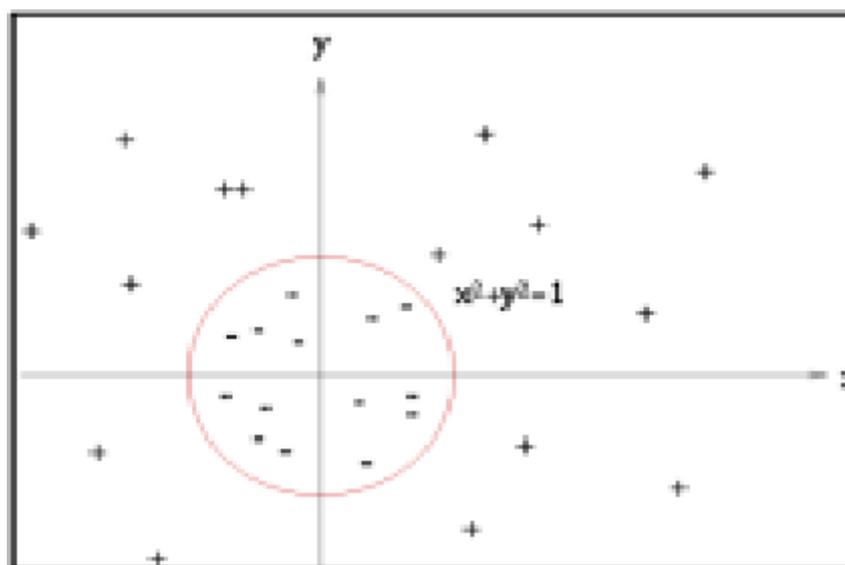


Figure 12 : classifications à deux classes avec une séparatrice non linéaire

➤ **La régression logistique :**

Mathématiquement, un modèle de régression logistique prédit la probabilité $P(Y)$ en fonction de X . C'est l'un des algorithmes les plus simples en apprentissage automatique et il est souvent utilisé pour résoudre divers problèmes de classification, tels que la détection de spam ou de maladies.

La régression logistique est une technique de classification supervisée qui vise à prédire la probabilité d'une variable cible. Elle analyse un ensemble de données où une ou plusieurs variables déterminent un résultat. Ce résultat est généralement mesuré avec une variable dichotomique, qui n'a que deux résultats possibles. L'objectif de la régression logistique est de trouver le meilleur modèle pour décrire la relation entre la variable dépendante (ou variable cible) et un ensemble de variables indépendantes (prédicteurs ou variables explicatives).

En règle générale, la régression logistique est utilisée pour des régressions binaires où les variables sont binaires, mais elle peut également être appliquée à d'autres catégories de variables cibles. En fonction du nombre de catégories, on peut classer la régression logistique en différents types.

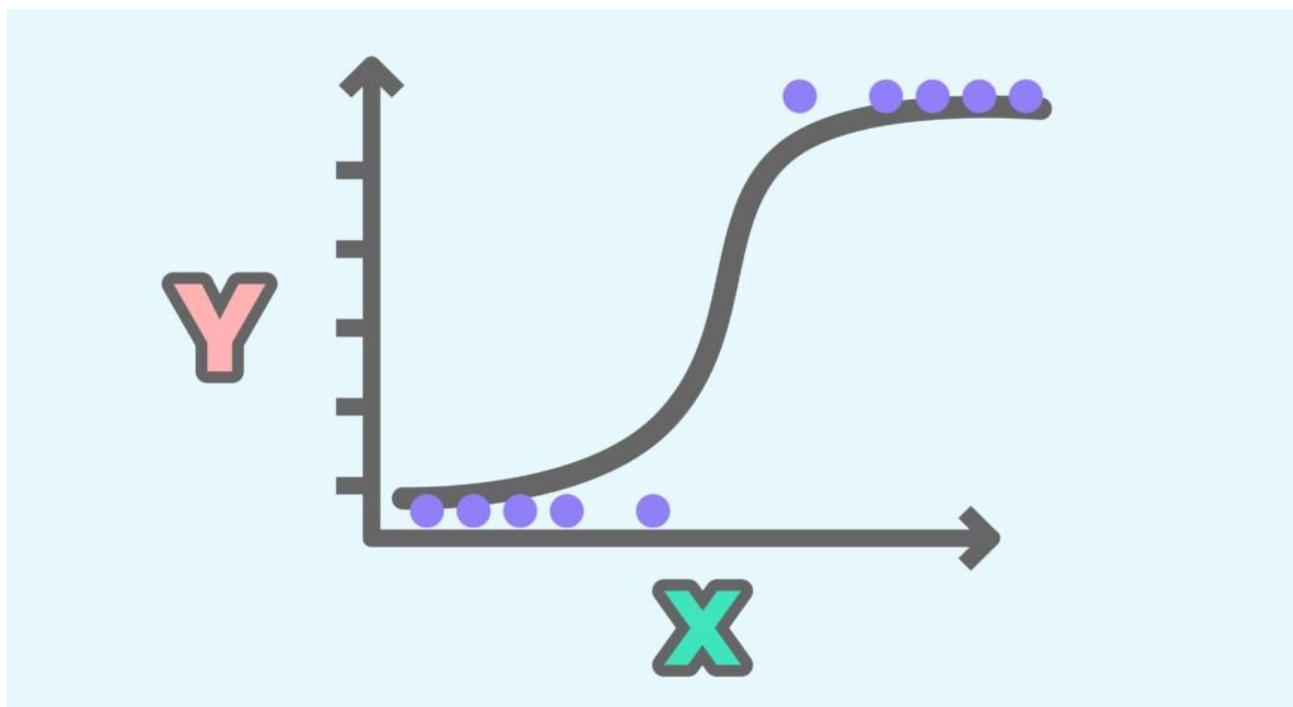


Figure 13 : régression logistique

➤ **Le modèle de mélange gaussien (GMM) :**

Le modèle de mélange gaussien (GMM) est une méthode d'apprentissage automatique probabiliste largement utilisée pour modéliser la distribution de données complexes. Contrairement à d'autres algorithmes de clustering qui assignent chaque observation à un seul cluster, le GMM considère chaque cluster comme une distribution gaussienne multidimensionnelle. Il suppose que les données sont générées à partir d'un mélange de plusieurs distributions gaussiennes, chacune étant associée à un cluster. L'objectif du GMM est d'ajuster les paramètres de ces distributions gaussiennes, tels que les moyennes et les covariances, afin de maximiser la vraisemblance des données observées. Pendant l'entraînement, le modèle cherche à estimer les poids de chaque distribution gaussienne, qui indiquent la probabilité qu'une observation appartienne à un cluster spécifique. Une fois le modèle entraîné, il peut être utilisé pour prédire les clusters auxquels de nouvelles observations appartiennent en calculant la probabilité de chaque observation appartenir à chaque cluster. En résumé, le GMM est un modèle flexible et puissant qui permet de capturer la structure complexe des données en utilisant des distributions gaussiennes.

Le modèle de mélange gaussien (GMM) se distingue par plusieurs aspects :

1. **Probabiliste** : Contrairement à certains algorithmes de clustering qui produisent des affectations de cluster déterministes, le GMM utilise des distributions de probabilité pour représenter chaque cluster, ce qui permet de modéliser l'incertitude quant à l'appartenance d'une observation à un cluster particulier.
2. **Souplesse** : Le GMM peut modéliser des données provenant de différentes distributions et de différentes formes, car il peut ajuster les paramètres de chaque distribution gaussienne de manière indépendante.
3. **Estimation des paramètres** : L'entraînement d'un GMM implique l'estimation des paramètres de chaque distribution gaussienne, y compris les moyennes, les covariances et les poids des composants, ce qui lui permet de s'adapter aux structures de données complexes.
4. **Complexité** : Le GMM peut être plus complexe à entraîner et à interpréter que certains algorithmes de clustering plus simples en raison du grand nombre de paramètres à estimer, notamment pour les données de haute dimensionnalité.
5. **Utilisations diverses** : Outre le clustering, le GMM est également utilisé pour la génération de données synthétiques, la détection d'anomalies et la modélisation des distributions de

probabilité dans des applications telles que la reconnaissance de formes et la vision par ordinateur.

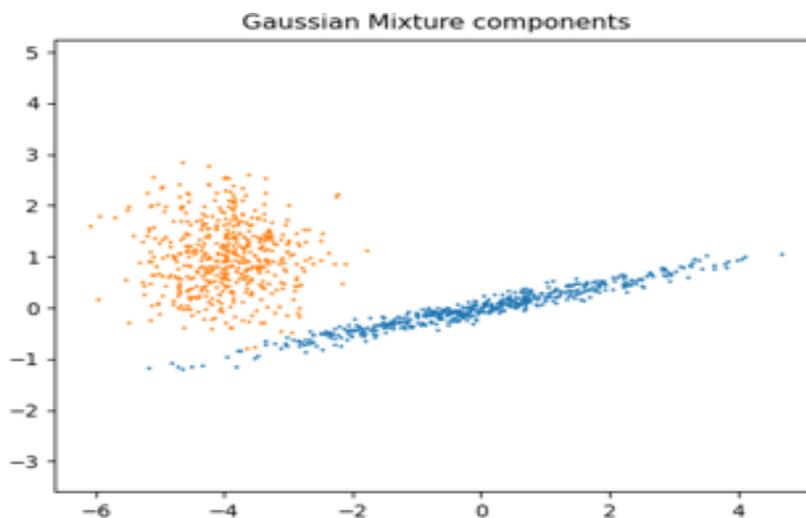


Figure 14 : Modèle de mélange gaussien (GMM)

D'autres modèles de clustering populaires, tels que le K-means et le DBSCAN, offrent des approches différentes pour regrouper les données. Le K-means divise les données en un nombre prédéfini de clusters en minimisant la distance moyenne entre chaque point de données et le centroïde de son cluster. Cependant, il nécessite une spécification préalable du nombre de clusters et peut être sensible aux points aberrants. En revanche, le DBSCAN identifie les clusters en se basant sur la densité des points de données : il trouve les régions denses de points dans l'espace des données, marquant les points comme noyau, frontière ou bruit. Bien qu'il puisse identifier des clusters de formes arbitraires et ne nécessite pas de spécification du nombre de clusters, il peut être sensible aux paramètres et à la densité des données. Ces deux approches offrent des alternatives valables pour le clustering, chacune avec ses propres forces et faiblesses selon le contexte d'application et les caractéristiques des données.

6. Apprentissage profond :

L'apprentissage profond est une sous-discipline de l'apprentissage automatique qui se concentre sur l'utilisation de réseaux de neurones artificiels avec plusieurs couches cachées pour modéliser et interpréter des données complexes. Contrairement à d'autres méthodes d'apprentissage automatique,

L'apprentissage profond est capable d'apprendre des représentations hiérarchiques des données à partir de grandes quantités de données non structurées, telles que des images, du texte ou du son. Cette approche permet au système d'apprendre des caractéristiques de haut niveau à partir de caractéristiques de bas niveau, ce qui lui permet de résoudre des tâches complexes telles que la reconnaissance d'objets dans des images, la traduction automatique et la génération de texte.

Bien que confier votre traitement à une machine puisse sembler une idée inhabituelle, l'apprentissage profond est en train de révolutionner le secteur de la santé. Le diagnostic médical, en particulier, bénéficie des capacités de l'apprentissage profond à analyser de grandes quantités de données médicales pour identifier rapidement et avec précision les maladies. Par exemple, les réseaux de neurones profonds peuvent être formés sur des ensembles de données d'imagerie médicale pour détecter les signes précurseurs de maladies telles que le cancer. De plus, l'utilisation de dispositifs portables et de capteurs connectés peut fournir des données en temps réel sur la santé des patients, permettant ainsi un suivi continu et personnalisé.

7. Conclusion :

En conclusion, l'apprentissage automatique représente une révolution dans le domaine de l'informatique et de l'intelligence artificielle. Grâce à cette approche, les machines peuvent non seulement exécuter des tâches spécifiques, mais aussi apprendre à partir des données pour améliorer leurs performances au fil du temps. De la reconnaissance vocale à la recommandation de produits en passant par la détection de fraudes, les applications de l'apprentissage automatique sont vastes et variées, et elles continuent d'évoluer rapidement.

Cependant, malgré ses avantages, l'apprentissage automatique n'est pas sans défis. La qualité des données, la sélection appropriée des algorithmes, et l'interprétabilité des modèles sont autant de considérations importantes à prendre en compte. De plus, il est essentiel de garantir l'équité et la transparence dans l'utilisation des modèles d'apprentissage automatique afin d'éviter les biais et les discriminations.

Dans l'ensemble, l'apprentissage automatique offre un potentiel énorme pour résoudre des problèmes complexes et apporter des innovations significatives dans de nombreux domaines. En continuant à investir dans la recherche et le développement de méthodes et d'outils plus avancés, nous pouvons exploiter pleinement les avantages de l'apprentissage automatique pour créer un avenir plus intelligent et plus prometteur.

Le apprentissage profond représente une avancée majeure dans le domaine de l'intelligence artificielle, offrant des possibilités extraordinaires pour résoudre des problèmes complexes de manière efficace et précise. Grâce à ses architectures de réseaux de neurones profonds et à ses algorithmes sophistiqués, le apprentissage profond permet aux machines d'apprendre des représentations hiérarchiques des données à partir de grandes quantités de données brutes. Cette capacité a ouvert la voie à des avancées significatives dans des domaines tels que la reconnaissance d'images, la traduction automatique, la génération de texte, la biologie computationnelle et bien d'autres encore.

Chapitre03

Méthodologie

1. Introduction :

Ce chapitre explique la méthode utilisée pour obtenir un ensemble de données basé sur les enregistrements vocaux dans le cadre de la détection de la maladie de Parkinson. Les données ont été collectées à partir d'échantillons vocaux de patients atteints de la maladie de Parkinson. Ensuite, différents algorithmes d'apprentissage automatique ont été appliqués pour entraîner des modèles capables de la maladie. Parmi ces algorithmes, Gaussien Mixture Model (GMM), Support vecteur Machine (SVM) et Régression logistique., puis les modèles ont été entraînés, testés et validés. Les performances des différents modèles ont été évaluées en termes de sensibilité, spécificité et précision pour déterminer leur capacité à détecter efficacement les signes de la maladie de Parkinson par voix.

2. Tâche :

L'architecture est expliquée comme suit : **(1)** acquisition de l'ensemble de données, **(2)** formation utilisant des algorithmes GMM, de régression logistique et SVM, et **(3)** configurations expérimentales.

- **Acquisition de l'ensemble de données** : Les données vocales des patients atteints de la maladie de Parkinson sont obtenues. Max A. Little, Patrick E. McSharry, Eric J. Hunter, Lorraine O. Ramig (2008), 'Aptitude des mesures de dysphonie pour la télésurveillance de la maladie de Parkinson', Transactions IEEE sur le génie biomédical (à paraître).
- **Formation utilisant des algorithmes GMM, de régression logistique et SVM** : Les modèles sont formés en utilisant les données étiquetées et préparées pour la formation à l'aide d'algorithmes GMM, de régression logistique et SVM.
- **Configurations expérimentales** : Les performances des modèles formés à l'aide des algorithmes sont évaluées, et le meilleur modèle pour la tâche de détection de la maladie de Parkinson est sélectionné."

3. Présentation de dataset :

Ce jeu de données est composé d'une gamme de mesures vocales biomédicales de 31 personnes, dont 23 atteintes de la maladie de Parkinson (MP). Chaque colonne dans le tableau est une mesure vocale particulière, et chaque ligne correspond à un enregistrement vocal parmi 195 de ces individus (colonne "Name"). L'objectif principal des données est de discriminer les personnes en bonne santé

de celles atteintes de la maladie de Parkinson, selon la colonne "statuts" qui est définie à 0 pour les personnes en bonne santé et à 1 pour les personnes atteintes de la maladie de Parkinson.

Les données sont au format CSV ASCII. Les lignes du fichier CSV contiennent une instance correspondant à un enregistrement vocal. Il y a environ six enregistrements par patient, le nom du patient est identifié dans la première colonne. Pour plus d'informations ou pour faire part de commentaires, veuillez contacter Max Little (littlem '@' robots.ox.ac.uk).

Des détails supplémentaires sont contenus dans la référence suivante - si vous utilisez ce jeu de données, veuillez le citer comme suit :

Max A. Little, Patrick E. McSharry, Eric J. Hunter, Lorraine O. Ramig (2008), 'Aptitude des mesures de dysphonie pour la télésurveillance de la maladie de Parkinson', Transactions IEEE sur le génie biomédical (à paraître).

Nous pouvons commencer à analyser les données et à expérimenter des algorithmes qui nous aideront à étudier détection de la maladie parkinson.

- La base de données utilisée est un fichier CSV
 - Nous avons 197 instances et les 9attributs suivants :
 - **Nom de patient.**
 - **MDVP:Fo** (Cette mesure représente la fréquence fondamentale de la voix en Hertz).
 - **MDVP:Fhi** (Elle représente la fréquence maximale de la voix en Hertz).
 - **MDVP:Flo** (est une mesure vocale dans le jeu de données. Elle représente la fréquence minimale de la voix en Hertz).
 - **MDVP:Jitter %** (Elle représente le pourcentage de variation de la fréquence de la voix, également connu sous le nom de jitter).
 - **MDVP:Jitter abs** (Il représente l'amplitude absolue des variations de la fréquence de la voix, également connue sous le nom de jitter absolu).
 - **MDVP:RAP** (Il représente le rapport d'amplitude du jitter, qui mesure les variations irrégulières de la fréquence de la voix, par rapport à la période.)
 - **MDVP:PPQ** (Il représente la perturbation du rapport de la période. C'est une mesure utilisée pour évaluer les variations de la fréquence de la voix).
 - **Jitter:DDP** (Il représente la déviation absolue de la différence de période, une mesure de l'irrégularité temporelle de la voix).

- **MDVP:Shimmer** (Shimmer est une mesure de l'instabilité de l'amplitude de la voix dans le jeu de données. Elle évalue la variation cyclique de l'amplitude du signal vocal, ce qui peut indiquer des irrégularités dans la voix).
- **MDVP: Shimmer** (est une mesure de l'instabilité de l'amplitude de la voix exprimée en décibels dans le jeu de données. Elle évalue la variation cyclique de l'amplitude du signal vocal, ce qui peut indiquer des irrégularités dans la voix).
- **Shimmer:APQ3** (est une mesure de la variabilité de l'amplitude de la voix calculée sur la base de la moyenne des différences entre les valeurs d'amplitude successives sur des intervalles de temps courts, spécifiquement sur les trois premiers cycles du signal. Cette mesure est utilisée pour évaluer la stabilité de la voix).
- **Shimmer:APQ5** (est une mesure de la variabilité de l'amplitude de la voix calculée sur la base de la moyenne des différences entre les valeurs d'amplitude successives sur des intervalles de temps courts, spécifiquement sur les cinq premiers cycles du signal. Cette mesure est utilisée pour évaluer la stabilité de la voix).
- **MDVP:APQ** (est une abréviation pour "Amplitude Perturbation Quotient", qui mesure la perturbation de l'amplitude de la voix. Il calcule la moyenne des différences entre les valeurs d'amplitude successives sur des intervalles de temps courts. Cette mesure est utilisée pour évaluer la stabilité de la voix et est particulièrement utile dans l'analyse des troubles de la voix comme la dysphonie).
- **Shimmer : DDA** (est une mesure de l'amplitude de la variation cyclique de l'intensité vocale. Il représente l'amplitude moyenne des différences entre les pics d'amplitude voisins dans un signal vocal, mesurée sur une période de temps spécifique. Cette mesure est souvent utilisée pour évaluer la stabilité et la régularité de la voix, en particulier dans le contexte des troubles vocaux tels que la dysphonie).
- **NHR (Noise-to- Harmonics Ratio)** : est un indicateur qui mesure le rapport entre le bruit et les harmoniques dans un signal vocal. Il est utilisé pour évaluer la présence de bruits indésirables dans la voix, tels que des tremblements ou des distorsions, par rapport aux harmoniques du signal vocal. Plus le NHR est élevé, plus le bruit est présent par rapport aux composantes harmoniques de la voix. Cette mesure est importante dans l'analyse des troubles de la voix, notamment dans le diagnostic de la dysphonie.
- **HNR : HNR (Harmonics-to-Noise Ratio)** est un indicateur qui mesure le rapport entre les harmoniques et le bruit dans un signal vocal. Il est utilisé pour évaluer la clarté et la qualité

du signal vocal, en mettant en évidence la présence d'harmoniques par rapport au niveau de bruit de fond. Un HNR élevé indique généralement un signal vocal de meilleure qualité, avec des harmoniques plus prononcées par rapport au bruit. Cette mesure est importante dans l'analyse vocale pour évaluer la présence de distorsions ou de bruits indésirables dans la voix.

- **"statuts"** : fait référence à l'état ou au statut des individus dans le contexte des données fournies. Dans le jeu de données mentionné, la colonne "statuts" indique si un individu est en bonne santé ou s'il est atteint de la maladie de Parkinson. Habituellement, dans de telles bases de données biomédicales, cette colonne contiendrait des valeurs binaires pour différencier les deux états, par exemple 0 pour indiquer une personne en bonne santé et 1 pour indiquer une personne atteinte de la maladie de Parkinson.
- **RPDE** : fait référence à l'Entropie relative au Dérive de Phase. C'est une mesure calculée à partir des enregistrements vocaux qui est utilisée pour caractériser la variabilité de la fréquence fondamentale de la voix. En termes simples, cela évalue la régularité des transitions entre les périodes de haute et basse fréquence dans la voix.
- **DFA** : désigne l'Analyse de la Fluctuation Détrendue (Detrended Fluctuation Analysis en anglais). C'est une méthode utilisée pour étudier la variabilité fractale dans les séries temporelles, y compris les signaux vocaux. Elle évalue la manière dont les fluctuations dans une série temporelle évoluent avec différentes échelles de temps, ce qui peut donner des informations sur la structure sous-jacente et la complexité du processus générant les données. Dans le contexte des enregistrements vocaux, DFA peut être utilisée pour étudier les modèles de variation dans la fréquence ou d'autres caractéristiques vocales à différentes échelles de temps.
- **spread1 ,spread2** : sont des mesures de la dispersion de la fréquence fondamentale.
- **D2** : fait référence mesures de complexité dynamique non linéaires.
- **PPE** : "PPE" peut signifier "Pitch Period Entropy", qui est une mesure de l'irrégularité des périodes de hauteur dans un signal vocal. C'est une caractéristique importante dans l'analyse des signaux vocaux et peut être utilisée pour évaluer certains aspects de la voix, tels que la stabilité et la régularité du pitch.

4. Les mesures de performances utilisées :

L'évaluation des performances du modèle d'apprentissage automatique est effectuée en générant prédite statuts et actuel statuts pour chaque algorithme de machine Learning à gagner aperçu du type d'erreur commise par l'apprentissage automatique modèle qui nous aide à comprendre les autres métriques telles que précision

- **Prédite statuts** : Ce sont des cas dans lesquels nous avons prédit oui (la personne est malade)
La valeur prédite correspond à la valeur réelle.
- **Actuel statuts** : Ce sont des cas dans lesquels nous avons prédit non (la personne n'est pas malade ou oui est malade), La valeur prédite correspond à la valeur réelle.
- **Parkinson détection** : Ce sont des cas dans lesquels nous avons comparé entre la prédite statuts et actuel statuts.
 - **Vrai** : correspondance des informations. (1_1)
la valeur prédite correspond à la valeur réelle.
 - **Faux** : n'est pas correspondance des informations. (1_0)
La valeur prédite a été faussement prédite.
- **Vrai pourcentage** : Le pourcentage de correspondance des informations
- **Faux pourcentage** : Le pourcentage de non-concordance des informations

5. Conclusion :

La méthodologie utilisée dans cette étude pour détecter la maladie de Parkinson à partir d'enregistrements vocaux a été décrite en détail. Les étapes comprenaient l'acquisition de données à partir d'échantillons vocaux de patients atteints de la maladie, l'application d'algorithmes d'apprentissage automatique tels que les modèles à mélange gaussien (GMM), les machines à vecteurs de support (SVM) et la régression logistique pour former des modèles de détection, et enfin l'évaluation des performances des modèles formés.

Les données utilisées dans cette étude ont été obtenues à partir d'un ensemble de mesures vocales biomédicales de patients, avec l'objectif principal de différencier les individus en bonne santé de ceux atteints de la maladie de Parkinson. Les données ont été présentées sous forme de tableau CSV contenant différentes mesures vocales ainsi que des informations sur le statut de chaque individu en termes de santé ou de maladie de Parkinson.

Différentes mesures de performance ont été utilisées pour évaluer les modèles d'apprentissage automatique, notamment en comparant les valeurs prédites avec les valeurs réelles pour chaque algorithme. Cela a permis de calculer les pourcentages de correspondance et de non-correspondance des informations, fournissant ainsi un aperçu du type d'erreurs commises par les modèles.

Chapitre04

Expérimentations et résultat

1. Introduction :

Nous pouvons commencer à appliquer des techniques d'apprentissage automatique pour la classification dans un ensemble de données qui décrit les malades de parkinson.

Compte tenu des données médicales que nous pouvons recueillir sur les personnes, nous devrions être en mesure de faire de meilleures prédictions sur la probabilité qu'une personne souffre de l'apparition du parkinson, et donc d'agir de manière appropriée pour aider.

Dans ce chapitre, nous décrivons les procédures suivies pendant la création du modèle de détection de la maladie parkinson incluant le jeu de données utilisé, scénario expérimental, résultats et explications. Différents algorithmes d'apprentissage automatique ont été utilisés sur l'ensemble de données parkinson ..La comparaison entre les résultats obtenu nous a permet de sélectionner un meilleur algorithme pour notre modèle.

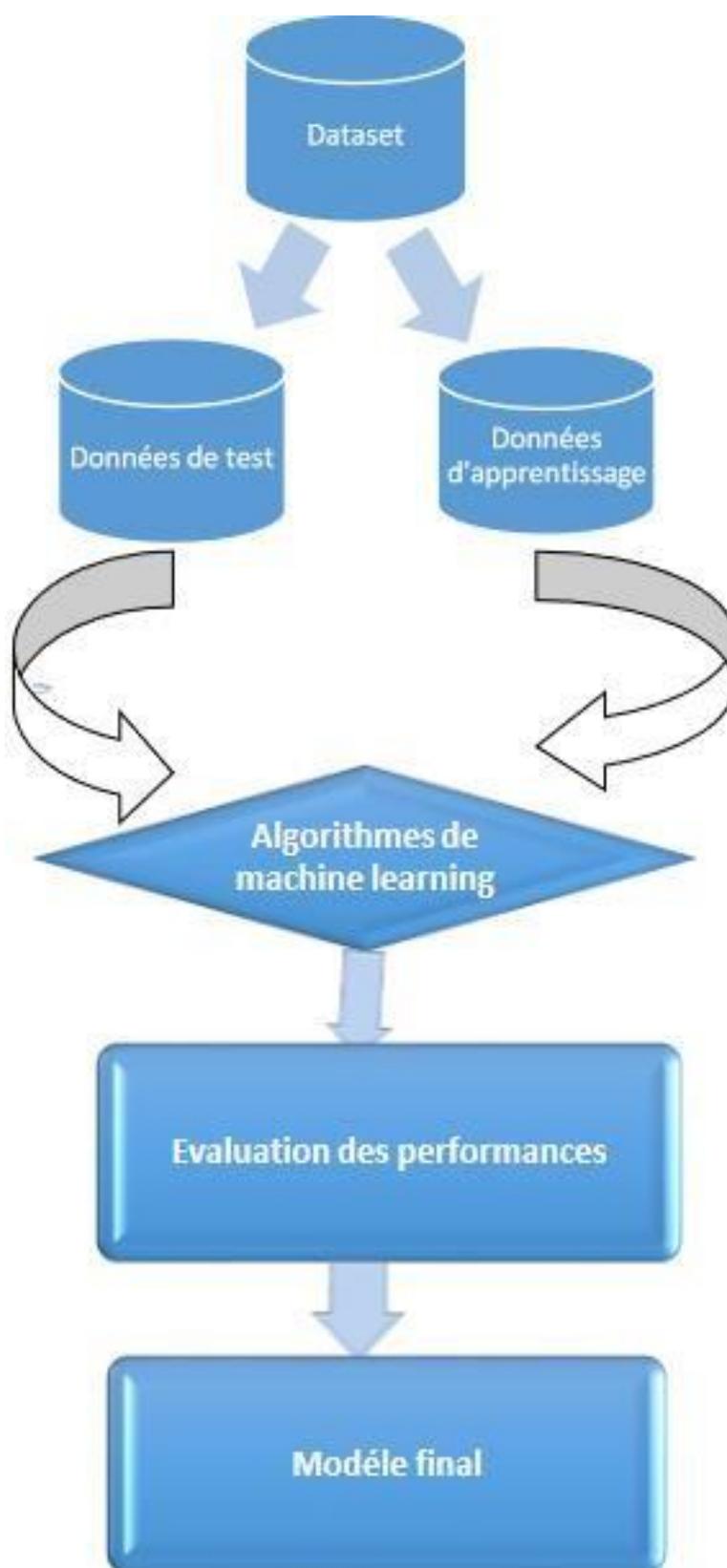


Figure 15 : le modèle proposé

2. Présentation des outils utilisés :

Le matériel joue un rôle essentiel dans les performances du modèle, Nous avons utilisé un ordinateur portable sous Windows 10, Système d'exploitation 64 bits avec 4Go de RAM. Le processeur est core i3 avec une vitesse d'horloge de 2,6GHz.

➤ Langage de programmation :

- **Python** : Python est un langage de programmation interprété, Développé en 1989. Il est utilisé pour de nombreuses applications différentes. Il est utilisé par des développeurs de logiciels professionnels dans des endroits tels que Google, la NASA. Ainsi python est le langage le plus utilisé dans le domaine d'apprentissage automatique. Ses principales caractéristiques sont :
 - « Open-source » : son utilisation est gratuite et les fichiers sources sont disponibles et modifiables, Simple et très lisible.
 - Doté d'une bibliothèque de base très fournie
 - Importante quantité de bibliothèques disponibles : pour le calcul scientifique, les statistiques, les bases de données, la visualisation.
 - Grande portabilité : indépendant vis à vis du système d'exploitation (linux, Windows, MacOS)



Figure 16 : logo python

- **Visual studio code**

Visual Studio Code est un éditeur de code source développé par Microsoft. Connue pour sa légèreté, sa polyvalence et son extensibilité, il est largement utilisé par les développeurs pour écrire et déboguer du code dans une variété de langages de programmation. Visual Studio Code offre une interface utilisateur intuitive, des fonctionnalités avancées telles que la coloration syntaxique, l'autocomplétions intelligente, la navigation dans le code et l'intégration avec des outils de

développement populaires tels que Git. Il est disponible sur plusieurs plateformes, y compris Windows, MacOS et Linux, et il est largement apprécié pour sa communauté active qui crée une multitude d'extensions pour personnaliser et améliorer l'expérience de développement.



Figure 17: Visual studio code

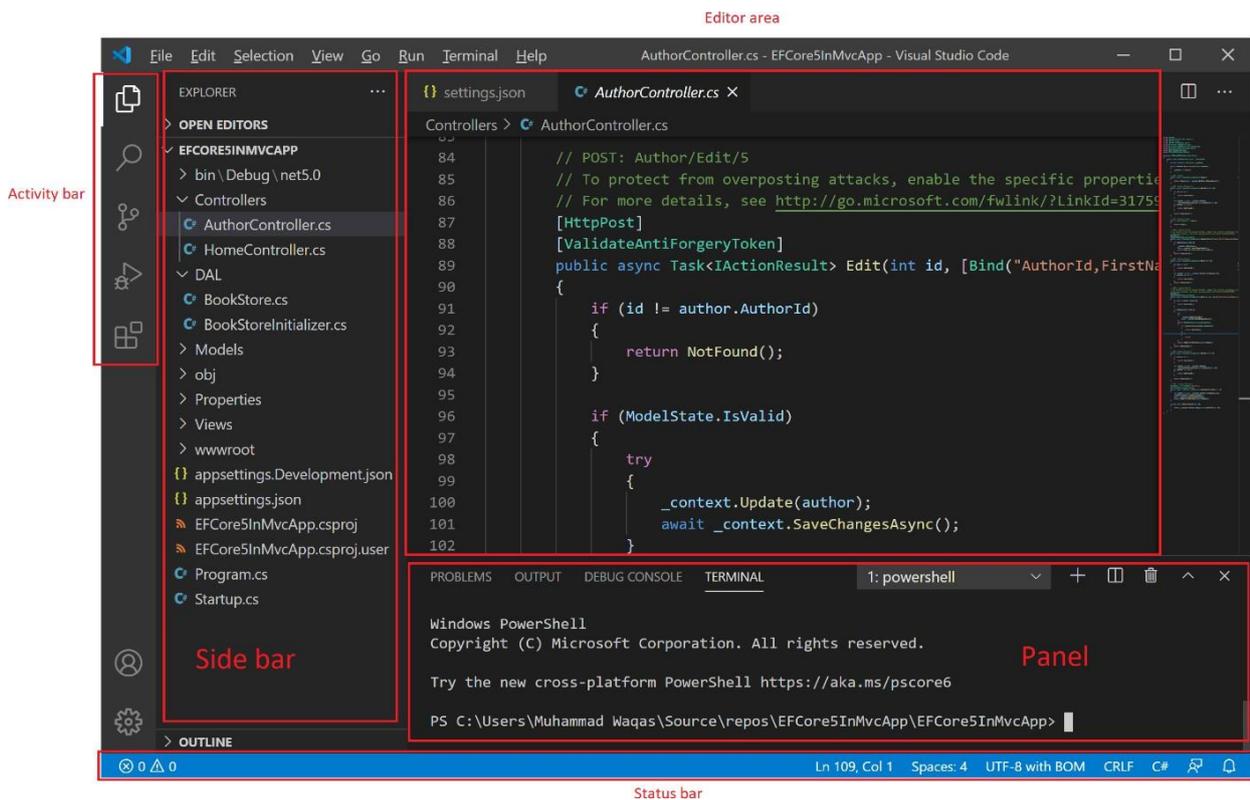


Figure 18: Interface Visual studio code .

➤ **Bibliothèques utilisées**

Pour traiter l'ensemble de données et mettre en œuvre l'apprentissage automatique, nous avons utilisé de nombreuses bibliothèques python.

- **Pandas** : la bibliothèque Pandas prend en charge l'analyse des données. Nous utilisons la bibliothèque pandas pour importer l'ensemble de données au format de fichier .CSV et pour manipuler les données
- **Matplotlib** : la bibliothèque Matplotlib est utilisée pour visualiser les données sous format graphique. Cette bibliothèque prend en charge le graphique à barres, le nuage de points et de nombreux autres graphiques qui aident à comprendre et analyser clairement les résultats obtenus.
- **Sklearn** : Scikit-learn est une bibliothèque libre Python destinée à l'apprentissage automatique. Elle comprend des fonctions pour estimer des forêts aléatoires, des régressions logistiques, des algorithmes de classification, et les machines à vecteurs de support. Elle est conçue pour s'harmoniser avec d'autres bibliothèques libres Python, notamment NumPy et SciPy. La bibliothèque Sklearn est principalement utilisée pour créer la matrice de confusion, pour diviser un ensemble, pour effectuer le prétraitement des données et pour la procédure d'ingénierie des fonctionnalités.
- **Numpy** : NumPy est une bibliothèque open source de calcul numérique en Python. Elle fournit des structures de données et des fonctions pour manipuler des tableaux et des matrices multidimensionnels, ainsi que des outils pour effectuer des opérations mathématiques et statistiques avancées. NumPy est largement utilisé dans le domaine de la science des données, de l'apprentissage automatique, de la recherche en physique, en bio-informatique et dans de nombreux autres domaines pour sa performance élevée et sa facilité d'utilisation. En résumé, NumPy est essentiel pour le calcul numérique en Python.

3. Expérimentations et discussions

Nous allons diviser notre ensemble de données en un échantillon d'entraînement de 80% qui permet à l'algorithme de s'entraîner dessus et un échantillon de test de 20% qui est inconnu pour l'algorithme, qui consiste à tester et à évaluer l'algorithme pour sa capacité de prédiction sur des données nouvelles, pour faire cela on utilise la fonction : `train_test_split`.

```
# Division du jeu de données en ensembles d'entraînement et de test
X_train, X_test, y_train, y_test = train_test_split(X_scaled, y, test_size=0.2, random_state=42) # Division des données
```

Figure 19 : Séparation de données entraînement et de test

- Nous avons appliqué, les différents algorithmes de machine Learning sur le dataset et nous avons obtenus les résultats suivants :

Ces images montre une interface utilisateur d'un programme d'analyse de la maladie de Parkinson a l'aide des algorithmes d'apprentissage, le programme utilise un ensemble de données de caractéristiques vocales pour prédire si un patient est atteint de la maladie de parkinson ou no .

L'image affiche un graphique et un tableau.

- **Le graphique :**

Le modèle a été entraîné sur un ensemble de données provenant de patients atteint et non atteint de la maladie de parkinson.

L'axe des x graphique représentent les performances du modèle sur les données d'entraînement et les données de test.

L'axe des y représente la probabilité que le patient soit atteint de la maladie de parkinson.

- **Le tableau :**

Le tableau répertorie les performances du modèle sur un ensemble des patients. Chaque ligne correspond à un patient et présente ces informations, parmi ces informations :

(le statut réel du patient « atteint ou no de la maladie de parkinson » , la probabilité prédite par le modèle que la patient atteint la maladie , la prédiction du modèle malade ou non , pourcentage de patients correctement classés par le modèle , et le pourcentage de patient mal

classés).

- **Régression logistique :**

- **La prédiction :**

Affiche la probabilité du modèle dans cet exemple, la probabilité qui le patient atteint de la maladie de parkinson est 85%

La courbe des données de test est n'est pas très proche de 1 ce qui signifie que le modèle est

capable de généraliser correctement à de nouvelles données.

Pourcentage de patients correctement classés : 84.61%

Pourcentage de patients mal classés : 15.38%

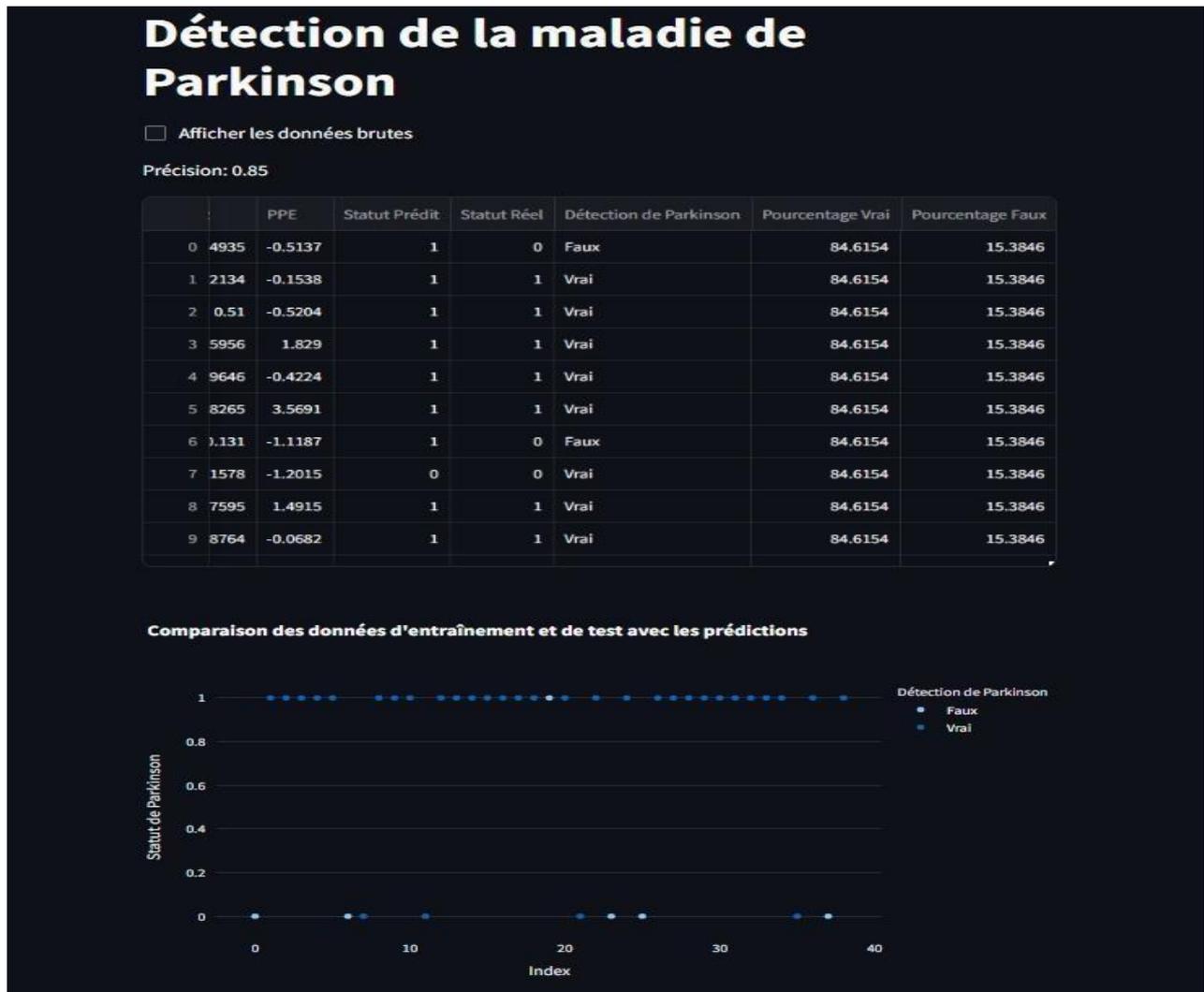


Figure 20 : résultats de comparaison des données d'entraînement et de test avec les prédictions régression logistique .

○ **Modèle de support vectoriel (SVM) :**

• **La prédiction :**

Affiche la probabilité du modèle dans cet exemple, la probabilité que le patient atteint de la maladie de parkinson est 90%

La courbe des données de test est également proche de 1 ce qui signifie que le modèle est capable de généraliser correctement à de nouvelles données.

Pourcentage de patients correctement classés : 89.74%

Pourcentage de patients mal classés : 10.25

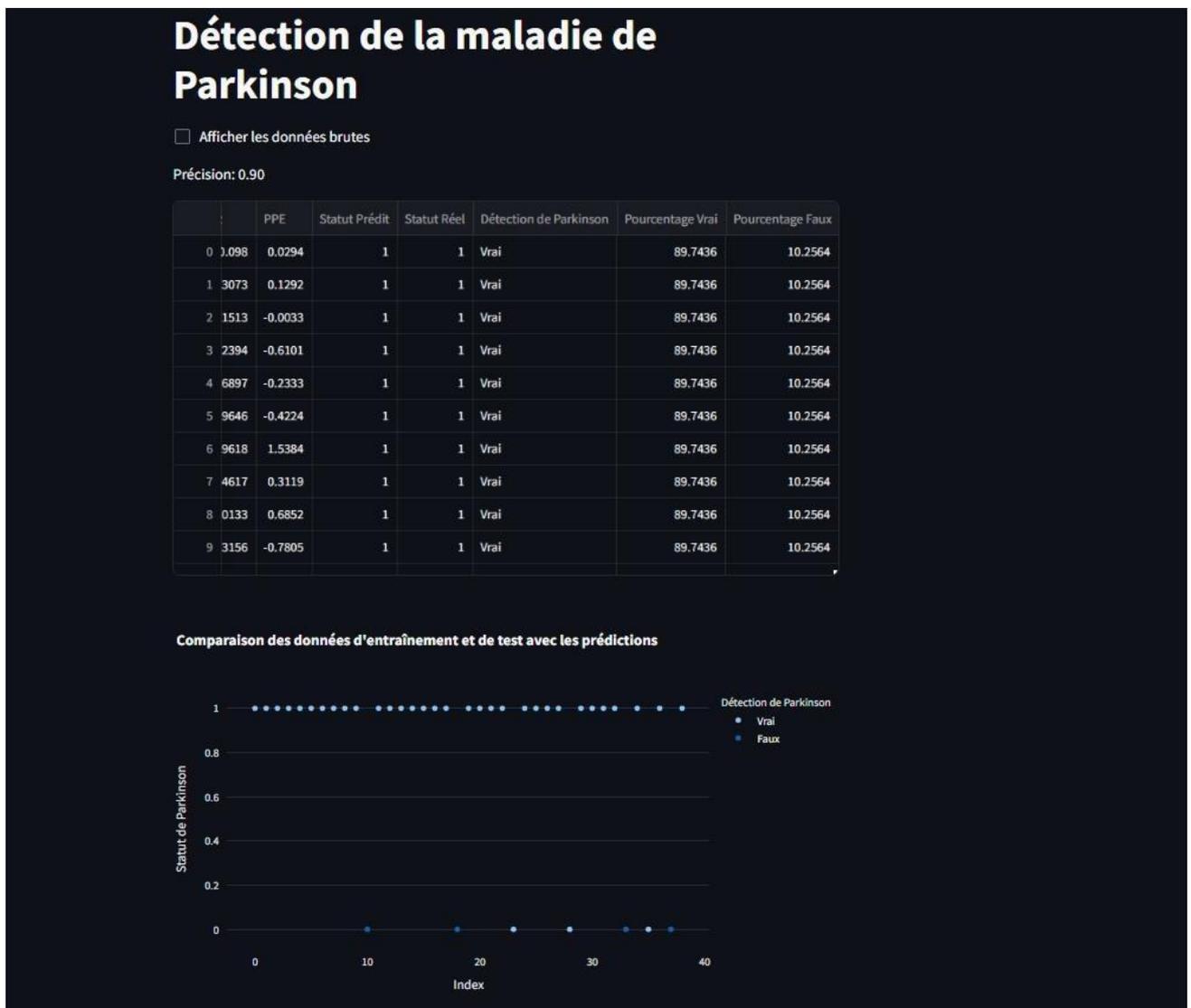


Figure 21 : résultats de comparaison des données d’entraînement et de test avec les prédictions SVM.

○ **Modèle de gaussien mixture (GMM) :**

Affiche la probabilité du modèle dans cet exemple, la probabilité qui le patient atteint de la maladie de parkinson est 61.54%

La courbe des données de test est également proche de 1 ce qui signifie que le modèle est capable de généraliser correctement à de nouvelles données

Pourcentage de patients correctement classés : 61.53%

Pourcentage de patients mal classés : 38.46%

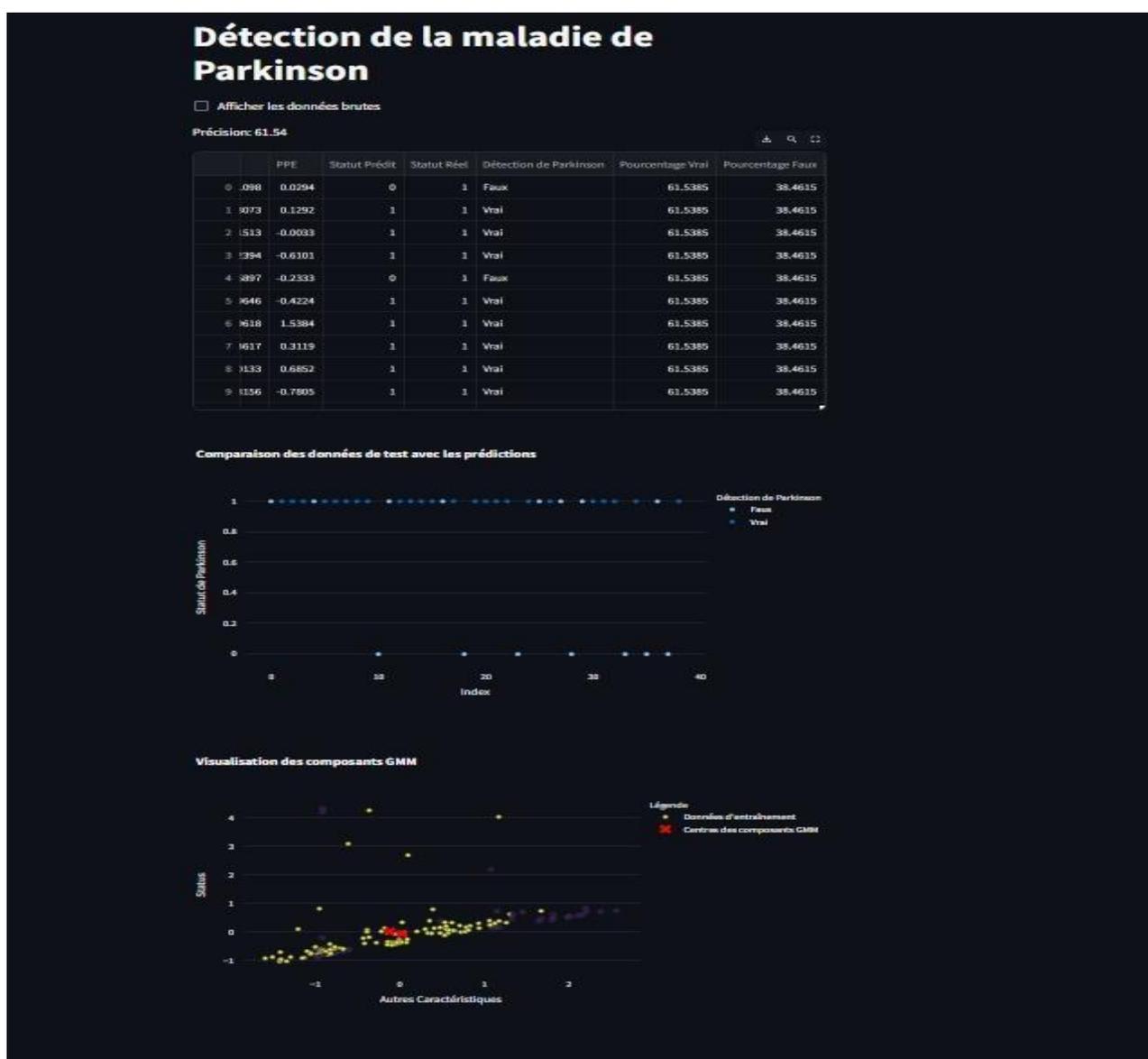


Figure 22 :résultats de comparaison des données d’entrainement et de test avec les prédictions GMM .

4. comparaisons :

Après avoir testé les algorithmes d'apprentissage automatique, on remarque que la plupart des algorithmes donnent des bons résultats. Ces comparaisons globales sur différentes métriques d'évaluation des algorithmes d'apprentissage automatique

Algorithme \ Métrique	GMM		SVM		REGRESSION LOGISTIQUE	
	0	1	0	1	0	1
Précision	38.46%	61.53%	10.25%	89.74%	15.38%	84.61%
Accuracy%	61.53%		90%		85%	

Tableau 1 : comparaison entre les algorithmes d'apprentissage .

L'algorithme le plus optimal qui répond le mieux à notre problématique c'est SVM qui nous a donné une meilleure accuracy de 90% comparé aux autres algorithmes testés : GMM 61.53 %, régression logistique 85% respectivement.

5. Conclusion :

Dans ce chapitre, nous avons présenté notre phase d'expérimentations et résultats, nous avons appliqué l'ensemble de 3 algorithmes d'apprentissage automatique sur un ensemble de données 'Parkinson ', et nous avons calculé les mesures qui nous permettent de choisir le meilleur algorithme.

Conclusion générale

CONCLUSION

En résumé, la maladie de Parkinson présente des défis significatifs en termes de détection précoce et de suivi de sa progression. Le développement de systèmes innovants, basés sur l'analyse des signaux vocaux et utilisant des techniques avancées d'apprentissage automatique et d'apprentissage en profondeur, pourrait révolutionner la manière dont cette maladie est diagnostiquée et gérée. Un tel système permettrait non seulement une détection plus rapide et plus précise de la maladie, mais aussi une surveillance continue de son évolution, améliorant ainsi la qualité de vie des patients. La construction d'une base de données robuste, la conception d'algorithmes performants et l'évaluation rigoureuse de ces outils sont des étapes essentielles pour atteindre cet objectif ambitieux.

Dans ce travail, nous avons proposé une approche pour la détection de la maladie de Parkinson, c'est-à-dire, déterminer si une personne est atteinte de la maladie de Parkinson ou non. En testant différents algorithmes de Machine Learning et en sélectionnant les algorithmes les plus optimaux pour notre problématique.

Afin d'aboutir à ces résultats, nous devons lire et étudier plusieurs publications et articles pour voir ce qui se fait de mieux dans le domaine de la détection et pour pouvoir concevoir notre propre approche. D'après les expériences rapportées dans les articles, nous remarquons qu'il est important de tester plusieurs algorithmes pour savoir lequel est le mieux adapté à une problématique donnée, car il n'y a pas de règle générale qui stipule qu'un algorithme est meilleur qu'un autre.

Grâce aux avancées de l'intelligence artificielle, comme le Machine Learning, nous pouvons traiter des problématiques plus complexes et obtenir de bons résultats, ce qui est crucial dans le domaine de la santé où la précision peut sauver des vies humaines. L'intelligence artificielle offre de nombreux avantages pour explorer le paysage de la santé et promet d'innover dans le système de santé pour un avenir meilleur.

Cette étude nous a permis de mieux comprendre les applications du Machine Learning dans le diagnostic médical. Ce fut également une leçon importante sur la conversion des données et l'analyse des algorithmes.

Il est quelque peu regrettable que de nombreux ensembles de données médicales soient petits (cela peut être dû à la confidentialité des patients), car un ensemble de données plus grand nous donnerait plus de flexibilité et de robustesse dans l'analyse. Cependant, nous croyons que cette étude est un bon début pour construire des moyens d'aider à diagnostiquer les patients, en comblant l'écart entre les médecins et les grands ensembles de données.

Références :

- [Braak and Tredici, 2017] Braak, H., & Tredici, K. D. (2017). Neuropathological Staging of Brain Pathology in Sporadic Parkinson's disease: Separating the Wheat from the Chaff. *Journal of Parkinsons Disease*, 7(s1). doi: 10.3233/jpd-179001 .
- [Kurlowicz and Wallace, 1999] Kurlowicz, L. & Wallace, M. (1999). The Mini-Mental State Examination (MMSE). *Journal of Gerontological Nursing*, 25(5), 8-9.
- [Perlmutter, 2009] Perlmutter, J. S. (2009). Assessment of Parkinson Disease Manifestations. *Current Protocols in Neuroscience*, 49(1). doi: 10.1002/0471142301.ns1001s49
- [Petersen, 2017] Petersen, M. (2017). Tractography and Neurosurgical Targeting in Deep Brain Stimulation for Parkinson's Disease. Aarhus University, Denmark
- [Port, 2019] Port, B. (2019, September 9). What brain areas are affected by Parkinson's? Retrieved from <https://medium.com/parkinsons-uk/what-brain-areas-are-affected-by-parkinsons-8c14dbf30954>.
- [Weiner et al., 2013] Weiner, W. J., Shulman, L. M., & Lang, A. E. (2013). *Parkinsons disease: a complete guide for patients and families*. Baltimore: The Johns Hopkins University Press.