



HAL
open science

Extraction des caractéristiques pour la classification de la maladie de Parkinson

Soumia Benikhlef, Bendimerad El Batoul, Nesma Settouti

► **To cite this version:**

Soumia Benikhlef, Bendimerad El Batoul, Nesma Settouti. Extraction des caractéristiques pour la classification de la maladie de Parkinson. 2013. hal-00846805

HAL Id: hal-00846805

<https://hal.science/hal-00846805>

Submitted on 20 Jul 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Extraction des caractéristiques pour la classification de la maladie de Parkinson.

Soumia Benikhlef*, El Batoul Bendimerad* et Nesma Settouti*

*Département Génie Électrique et Électronique,
Laboratoire Génie Biomédical
Université Abou Bekr Belkaid – Tlemcen,
B.P.230- Tlemcen 13000, Algérie

Table des matières

I	Contexte Médicale	3
1	Les notions médicales sur parkinson	3
1.1	Définition de la maladie	3
1.2	Symptômes	3
1.3	Les causes	4
1.4	L'évolution de la maladie	4
1.5	Diagnostic	4
1.5.1	Étape d'observation :	5
1.5.2	Etape d'examen :	5
1.6	Traitement	5
1.7	Statistiques	5
2	Les dernières recherches sur la maladie	6
2.1	Pourquoi la voix ?	6
2.2	D'où est venue l'idée ?	6
II	Expérimentations	7
3	Présentation de la base de données	7
3.1	Ensemble de données d'information	7
3.2	Définition des paramètres	7
3.3	Explication des paramètres	8
III	Application et interprétation	8

4	La réduction des paramètres	9
5	Analyse en composantes indépendantes	9
5.1	Le choix de l'Analyse en composantes indépendantes "ACI"	9
5.2	Etat de l'art	10
5.3	Définition	10
5.4	Modèle et objectif	10
5.5	Les avantages de l'ACI	11
5.6	Algorithme FastICA	11
6	Classification par réseau de neurones	11
6.1	Le choix du réseau de neurones	12
6.2	Les réseaux de neurones artificiels	12
6.3	L'algorithme d'apprentissage	13
6.4	Les avantages et les inconvénients des RNs	14
7	Application	14
7.1	Classification sans réduction	14
7.2	Classification avec réduction (ACI)	14
8	Discussion des résultats	16
9	Conclusion	17

Table des figures

1	Schéma représentatif de la dopamine au sein d'un neurone[Vas40]	4
2	Les 3 types de traitements pour soulager les malades [Dep12]	5
3	Modèle générale de la représentation de données	10
4	Modèle général d'un réseau de neurone artificiel	12
5	Réduction à 3 Composantes indépendantes	15
6	Réduction à 5 Composantes indépendantes	15
7	Réduction à 10 Composantes indépendantes	15
8	Réduction à 15 Composantes indépendantes	16

Liste des tableaux

1	Résumé de la base [Lit08]	7
2	Définition des paramètres	8
3	Avantages et inconvénients des RNs	14
4	Performances de classification sans réduction	14
5	Performances de classification avec réduction	16
6	Tableau des paramètres pertinents	17
7	Performances de classification avec et sans réduction	17

Introduction

Un diagnostic médical est le résultat du raisonnement d'un médecin avec une décision très souvent pris à partir de plusieurs informations.

Pour assurer un diagnostic exacte un système d'aide au diagnostic aide le médecin à réduire au minimum les erreurs possibles qui peuvent survenir pendant le diagnostic d'une maladie et classer les patients suivant deux catégories 'malade' et 'non malade' en analysant les différentes valeurs des paramètres pris en considération pour le dépistage de cette maladie. Des techniques telles que les réseaux de neurones sont le plus couramment utilisée dans les systèmes de classification et sont développés par le grand nombres d'équipe de recherches mais avec la présence d'attributs redondants ou attributs fortement bruités dans les bases de données, la performance du système peut se déclinier. ce qui nécessite l'utilisation des techniques de sélection de variables qui ont pour objet de sélectionner ou d'extraire un sous ensemble optimal de caractéristiques ou de paramètres les plus pertinents pour faire un meilleur apprentissage et assurer une bonne performance du système de classification.

Première partie

Contexte Médicale

1 Les notions médicales sur parkinson

1.1 Définition de la maladie

- La maladie de Parkinson est une maladie neurologique chronique dégénérative (perte progressive des neurones) affectant le système nerveux central responsable de troubles essentiellement moteurs d'évolution progressive.
- Bien que l'âge moyen d'apparition de la maladie de Parkinson soit 57 ans, il arrive que la maladie débute pendant l'enfance et semble héréditaire. La maladie de Parkinson touche les hommes et les femmes.
- Même si elle n'est pas mortelle en soi, à son dernier stade, elle comporte des risques de pneumonie, de dépression grave et de décès.

1.2 Symptômes

- Tremblements des membres : elle touche surtout la main, les pieds et le menton.
- Akinésie ou bradykinésie : Impossibilité ou difficulté de faire des mouvements d'origine neurologique et non musculaire. Elle affecte également l'élocution et le "langage corporel".
- Rigidité musculaire : une hypertonie musculaire touchant à la fois le rachis et les membres.
- D'autres symptômes peuvent s'ajouter : amaigrissement, constipation, hyper salivation, ralentissement de la déglutition, troubles du sommeil, de la parole et de l'écriture, crampes et Une incontinence urinaire.

1.3 Les causes

La maladie de Parkinson est une maladie qui correspond à la perte de neurones dans des zones noires précises du cerveau (Figure 1).

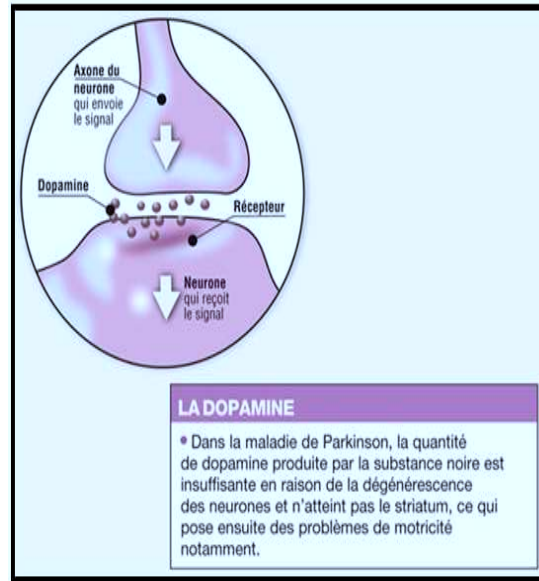


FIGURE 1 – Schéma représentatif de la dopamine au sein d'un neurone[Vas40]

Les neurones de la substance noire fabriquent normalement un message chimique(neurotransmetteur) appelé dopamine. Le manque de dopamine est responsable des signes caractéristiques de la maladie de parkinson

1.4 L'évolution de la maladie

On décrit trois étapes évolutives de la maladie de Parkinson :

- La "lune de miel" qui est une période variant de 3 à 8 ans et qui se définit par une vie pratiquement normale. C'est une période de meilleure efficacité du traitement.
- Les classiques fluctuations qui affectent 60% des patients après six ans d'évolution : akinésie. [Dep12]
- La 3ème période, la plus handicapante, correspond à la perte d'efficacité de la L-Dopa.

1.5 Diagnostic

Le diagnostic se repose sur deux étapes :

1.5.1 Étape d'observation :

Après un interrogatoire précis sur les symptômes constatés, le médecin observe la démarche et les tremblements éventuels. Il évalue aussi le langage, les mouvements des yeux et les réflexes.

1.5.2 Etape d'examen :

- Une prise de sang, chez un sujet appartenant, par exemple, à une famille de parkinsoniens. Cela peut aider à découvrir certains marqueurs biologiques atteints, aussi un taux particulièrement élevé de la protéine dans le sang, alpha-synucléine.
- Une échographie du crâne À l'échographie, le médecin peut repérer une modification de la substance noire du cerveau, ce qui peut être un indice.
- Une scintigraphie ou un IRM qui peut repérer les transporteurs membranaires de la dopamine.

1.6 Traitement

Il n'existe aucun traitement curatif L-DOPA, une molécule qui se transforme en dopamine dans l'organisme.

La prise de médicaments doit être répétée tout au long de la journée, pour maintenir un taux stable dans le sang. Malheureusement l'organisme s'habitue aux produits, ce qui rend le médicament moins efficace, au cours du temps (Figure 2).

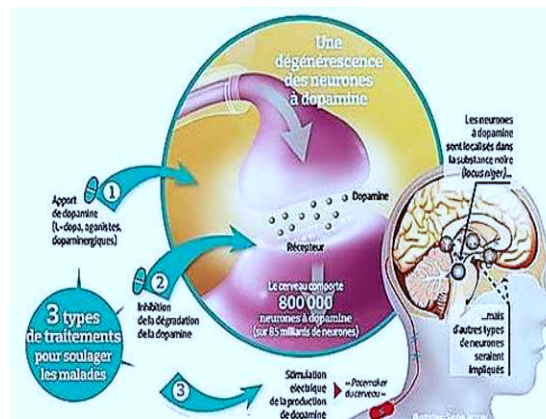


FIGURE 2 – Les 3 types de traitements pour soulager les malades [Dep12]

Il existe aussi les traitements chirurgicaux qu'il s'agit d'implanter des électrodes de stimulation électrique dans une région précise du cerveau ces électrodes envoient des impulsions électriques en profondeur, ce qui semble rétablir un fonctionnement normal du système.

1.7 Statistiques

La prévalence de la maladie de Parkinson en Europe est d'environ 1 personne pour 1000 habitants. 14.000 personnes déclarent chaque année cette maladie en France USA montrent qu'on

dénombrer 160 personnes atteintes de la maladie de Parkinson pour 100'000 habitants. Soit une prévalence d'exactly 1.6%. 6,5 millions de personnes sont touchées ans monde. Elle est plus fréquente chez les personnes âgées, atteignant plus 4% chez les personnes de plus de 85 ans [Vas40].

2 Les dernières recherches sur la maladie

Les recherches actuelles sont ciblées sur la détection de parkinson par la voix et sur cette idée se repose notre projet.

2.1 Pourquoi la voix ?

La voix est le résultat d'une coordination du larynx, du diaphragme, des cordes vocales, de la langue et des lèvres. Chez un malade de Parkinson, cette coordination est altérée. La voix a des rigidités, des faiblesses et des tremblements. Elle est un bon marqueur de la maladie. Il est possible qu'elle soit même l'une des premières fonctionnalités affectées. Je n'en suis pas encore sûr, mais, en tous cas, Parkinson se détecte par la voix, même à un stade très précoce.

2.2 D'où est venue l'idée ?

Une nouvelle piste vient d'être étudiée par un mathématicien du MIT (l'Institut technologique du Massachusetts), Max Little, qui s'est associé à deux autres chercheurs de l'université britannique d'Oxford. Son idée [TLMR10] a été de mettre au point un logiciel qui permet de détecter aisément cette maladie en analysant la voix du patient.

Le logiciel serait capable de déceler les infimes modifications de la voix qui seraient l'un des premiers signes de cette pathologie. Ce système permettrait d'ailleurs une analyse par simple appel téléphonique : l'idée est faire prononcer un «aaaaah» le plus long possible ce qui pourrait améliorer le suivi des patients, par exemple pour vérifier les résultats d'un ajustement de traitement. Pour développer ce système, le chercheur a rassemblé des échantillons de voix de 50 personnes atteintes de cette maladie, enregistrées chaque mois pendant 6 mois.

Il a également intégré des échantillons de voix de personnes en bonne santé, mais également de personnes présentant d'autres pathologies pouvant elles aussi altérer le son des cordes vocales. Il a ensuite développé un algorithme capable de détecter les particularités de ces échantillons, et de déceler les changements purement associés à la maladie. Le système, encore perfectible, sait aujourd'hui identifier les voix des personnes atteintes de la maladie de Parkinson avec un taux de réussite de 86%.

Le logiciel "apprend" ainsi à détecter les différences dans les modes vocaux. Grâce à ces premiers échantillons, il a appris à faire la différence entre les particularités engendrées par cette maladie et d'autres éventuels tremblements ou altérations dus à divers facteurs (rhume, tabagisme, chirurgie de la gorge...). Reste à analyser de nombreux échantillons complémentaires, d'une population la plus large possible afin de rendre ce système le plus perfectible possible.

Deuxième partie

Expérimentations

Introduction

Dans cette partie expérimentations nous avons fait appel au jeu de données crée par Max Little [Lit08], de l'Université d'Oxford, en collaboration avec le Centre national pour la voix et de la parole, Denver, Colorado, qui a enregistré les signaux de parole. L'étude originale a publié les méthodes d'extraction de caractéristiques pour les troubles de la voix en général.

3 Présentation de la base de données

3.1 Ensemble de données d'information

Cette base est composée d'une série de mesures vocales biomédicales de 31 personnes, 23 à la maladie de Parkinson (MP) [Lit08]. Chaque colonne de la table est une mesure voix particulière, et chaque ligne correspond un enregistrement vocal de 195 de ces personnes (colonne «name»).

L'objectif principal de ces données est de discriminer les personnes en bonne santé de ceux avec malade, selon colonne "statuts" qui sont mis à 0 pour la santé et 1 pour malade.

Les données sont au format CSV ASCII. Les lignes du fichier CSV contient une instance correspondant à un enregistrement de la voix. Il y a environ six enregistrements par patient.

Caractéristiques des données d'ensemble :	Multi variée	Nombre d'instances :	195	Zone : Vie
Caractéristiques d'attribut :	Réel	Nombre d'attribut	23	Date de Don 26/06/2008
Tâches associées :	classification	Valeurs manquantes ?	N/A	Nombre de visites du site Web : 53703

TABLE 1 – Résumé de la base [Lit08]

3.2 Définition des paramètres

Entrées de la colonne de la matrice (attributs) : Nom - Nom de l'objet ASCII et le numéro de l'enregistrement

PARAMÈTRES	DÉFINITION
MDVP : Fo (Hz)	Moyenne vocal fréquence fondamentale
MDVP : FHI (Hz)	Maximum vocal fréquence fondamentale
MDVP : Flo (Hz)	Minimum vocal fréquence fondamentale
MDVP : Jitter (%), MDVP : Jitter (Abs), MDVP : RAP, MDVP : PPQ, Jitter : DDP	Plusieurs mesures de variation de fréquence fondamentale
MDVP : Shimmer, MDVP : Shimmer (dB), Shimmer : apq3, Shimmer : APQ5, MDVP : APQ, Shimmer : DDA -	Plusieurs mesures de variation d'amplitude
NHR, HNR	Deux mesures de rapport de bruit à composantes tonales de la voix statut L'état de santé du sujet (un) - de Parkinson, (zéro) - saine
RPDE, D2	Deux dynamiques non linéaires mesures de complexité
DFA	Signal d'échelle fractale exposant
Spread1, spread2, PPE	Trois mesures non linéaires de variation de fréquence fondamentale

TABLE 2 – Définition des paramètres

3.3 Explication des paramètres

Fréquence fondamentale : La fréquence F0 moyenne apporte une mesure globale de la hauteur de la voix (aiguë, grave ...).

Jitter : la perturbation à court terme de la fréquence fondamentale du cycle vibratoire mesurer d'un cycle à l'autre (autrement dit cycle à cycle).

Shimmer : la mesure de stabilité de l'intensité, la perturbation a court terme de l'amplitude du cycle vibratoire mesurer d'un cycle à cycle.

DFA :« detrended fluctuation analysis » est une mesure qui quantifie la présence ou l'absence de propriété de corrélation.

Le rapport signal-bruit : Deux mesures permettant de quantifier le rapport entre l'énergie du bruit et l'énergie harmonique du signal : le rapport signal/bruit et l'énergie normalisée du bruit.

Troisième partie

Application et interprétation

Introduction

Nous nous intéressons dans ce rapport de recherche à une nouvelle piste qui peut fournir un diagnostic précoce de la maladie, l'idée est de mettre au point un logiciel qui permet de détecter aisément cette maladie c'est-à-dire qui consiste à classer le patient suivant deux situations « parkinsonien ou sain » en analysant un certain nombre de paramètres caractérisant la voix du patient.

Notre travail consiste à améliorer son taux de classification en changeant la technique de caractérisation et la technique de classification par rapport à ceux réalisés dans la littérature notamment nous citons les travaux principaux de Max [TLMR10],[TLM⁺12].

Dans ce rapport nous avons utilisé la même base de données médicale pour la conception du classifieur réseau de neurone et nous avons comparé les résultats obtenus avec le même classifieur tout en appliquant les résultats d'une étape de réduction des paramètres dans laquelle nous avons adopté la méthode de l'analyse en composantes indépendantes.

4 La réduction des paramètres

Le principe du traitement de données peut se résumer dans l'expression de "réduction de données". En effet, il s'agit généralement d'extraire, d'un ensemble important de données, l'information pertinente, qui pourra souvent se réduire à une réponse simple permet de simplifier l'analyse des données et obtenir plus de robustesse par rapport à la variabilité des données.

La réduction de dimensionnalité peut servir à :

- Compression des données.
- Visualisation des données en 2D ou 3D.
- Extraction des caractéristiques : potentiellement+fondamentales+explicatives+ compactes.
- Éliminer les redondances.
- Peut considérer comme pré traitement c à d meilleures représentations de départ pour un autre algorithme (classification ou régression)

5 Analyse en composantes indépendantes

5.1 Le choix de l'Analyse en composantes indépendantes "ACI"

L'ACI est souvent appliquée pour éliminer une grande variété d'artefacts des signaux, le but est de récupérer les sources indépendantes à partir d'observations qui sont des mélanges linéaires inconnues de sources indépendantes non observées, elle est très efficace pour réaliser la séparation

de source lorsque :

- Le milieu de mélange est linéaire et les retards de propagation sont négligeables.
- Les sources sont indépendantes.
- Le nombre de sources est le même que le nombre de capteurs.

Pour notre cas le système nerveux central traite des signaux multi capteurs dont chaque composant est un mélange inconnu de sources inconnues supposées indépendantes.

5.2 Etat de l'art

La soirée cocktail « cocktail party problem » est le problème type que tente de résoudre l'analyse en composantes indépendantes ACI. Lors d'une réception animée il s'agit d'isoler ce que dit chaque participant à partir d'enregistrements des voix ambiant [Car02]. Ce problème rentre dans le cadre plus général de la séparation aveugle de sources (ignorer la façon dont les sources se mélangent) qui consiste à retrouver un certain nombre de source à partir des observations d'un mélange de celles-ci.

5.3 Définition

Analyse en composantes indépendantes (ACI) est une technique statistique et informatique pour révéler les facteurs cachés qui sous-tendent ensembles de variables aléatoires, des mesures, ou des signaux.

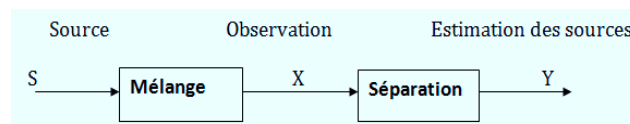


FIGURE 3 – Modèle générale de la représentation de données

ACI définit un modèle génératif pour les données observées à plusieurs variables, ce qui est généralement donné comme une grande base de données d'échantillons. Dans le modèle, les variables de données sont supposées être des mélanges linéaires de certaines variables latentes inconnus, et le système de mélange est également inconnu. Les variables latentes sont supposées non gaussien et mutuellement indépendants, et ils sont appelés les composantes indépendantes des données observées. Ces composants indépendants, aussi appelé sources ou facteurs, peut être trouvée par l'ACI.

5.4 Modèle et objectif

L'ACI cherche à expliquer un vecteur aléatoire n-dimensionnel en termes de composantes qui soient linéairement et statistiquement indépendantes. Autrement dit :

$$x(n) = A.s(n)$$

Où :

A est une matrice de mélange de taille $N \times N$

$s(n)$ et $x(n)$ sont respectivement les vecteurs de taille N des sources et des observations.

Cette méthode doit vérifier ces deux conditions [HO00] :

- Suivre une distribution normale (gaussienne),
- Le rang de la matrice A doit être égal au nombre de sources (à retrouver).

5.5 Les avantages de l'ACI

- Une technique puissante capable de trouver les facteurs fondamentaux quand les méthodes classiques échouent complètement
- Elle est l'une des approches utilisées pour la séparation aveugle des sources.
- L'ACP cherche à extraire des variables corrélées, en se limitant à imposer la contrainte d'indépendance aux statistiques d'ordre deux (matrice de covariance), l'ACI en revanche, cherche l'indépendance des statistiques d'ordre supérieur à deux des variables mesurées.
- L'ACI cherche la minimisation de la corrélation entre les composantes tout en maximisant l'aspect non gaussien de chaque composante.
- Elle cherche à rendre les composantes le plus indépendamment possible au sens de maximisé la fonction de contraste.
- Dans le cadre de l'ACI, une fonction de contraste mesure la dépendance statistique entre les composantes.
- Les données analysées par ACI ont pu provenir de beaucoup et de différents genres de champs d'application.

5.6 Algorithme FastICA

Une mise en œuvre particulièrement utile et largement utilisée de l'ACI est l'algorithme FastICA ce qui est très rapide et robuste. Il fonctionne bien avec de grands ensembles de données, et même dans des conditions un peu bruyantes. FastICA utilise un schéma d'optimisation de point fixe sur la base de Newton-itération et une fonction d'objectif liée à la néguentropie. L'algorithme FastICA peut rechercher les composants indépendants un à la fois ou tous à la fois. Sa performance peut aussi être à l'écoute un peu en choisissant parmi une gamme de non-linéarités.

6 Classification par réseau de neurones

Introduction

Les techniques de classification visent toutes à répartir n individus, caractérisés par p variables X_1, X_2, \dots, X_p en un certain nombre m de sous-groupes aussi homogènes que possible.

Ces techniques consistent à élaborer des règles de classification à partir des connaissances disponibles a priori ; il s'agit de la phase d'apprentissage via des algorithmes d'apprentissage. Le but de ces algorithmes est de produire des règles de classification afin de prédire la classe d'affectation d'un nouveau cas.

6.1 Le choix du réseau de neurones

Les réseaux de neurones se montrent très efficaces particulièrement pour des problèmes de classification, et surtout de prédiction. Leur principale caractéristique est l'aspect non linéaire de leur réponse, et ils sont ainsi une alternative prometteuse aux techniques plus traditionnelles pour des tâches de prédiction temporelle non linéaires.

6.2 Les réseaux de neurones artificiels

Définition :

Le réseau de neurone artificiel est un modèle de calcul dont la conception est très schématiquement inspirée du fonctionnement des neurones biologiques. Ils sont généralement optimisés par des méthodes d'apprentissage de type probabiliste, en particulier bayésien. Un réseau de neurones est en général composé d'une succession de couches dont chacune prend ses entrées sur les sorties de la précédente. Chaque couche (i) est composée de N_i neurones, prenant leurs entrées sur les N_{i-1} neurones de la couche précédente [TdAL09]. À chaque synapse est associé un poids synaptique, de sorte que les N_{i-1} sont multipliés par ce poids, puis additionnés par les neurones de niveau i , ce qui est équivalent à multiplier le vecteur d'entrée par une matrice de Transformation.

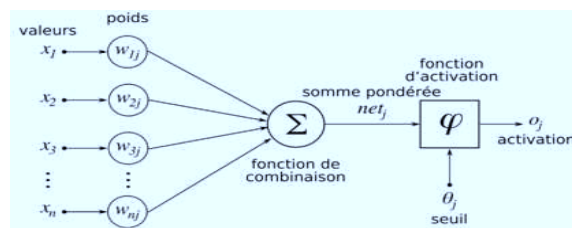


FIGURE 4 – Modèle général d'un réseau de neurone artificiel

Perceptron monocouche :

Un perceptron est un modèle pour collecter, acquérir ou percevoir des données d'entrées (des images, des textes etc.) ensuite un traitement est fait pour prendre une décision.

Un perceptron monocouche est composé d'une seule couche de connexion (poids synaptique w_i) entre l'entrée et la sortie. Pour un perceptron monocouche il y a deux tâches à réaliser

- La première tâche : déterminer le nombre de neurones et le nombre de connexion entre les neurones (chercher la structure du perceptron).
- La deuxième tâche : chercher et ajuster les valeurs des poids synaptiques w_i et les seuils (calculer les paramètres du perceptron).

Perceptron multicouche :

Un perceptron multicouche est composé d'une couche d'entrée et d'une couche de sortie en plus

d'une ou plusieurs couches cachés (composée de plusieurs neurones cachés. Nous menons deux types d'apprentissages :

- Apprentissage paramétrique : calcule les valeurs des poids synaptiques et les seuils via un algorithme d'apprentissage.
- Apprentissage structurel : on fixe la structure du réseau par le calcul de nombre de couche caché et le nombre de neurone dans toutes les couches et en même temps le nombre des connexions entre les neurones.

Le réseau de neurone multicouche est basé sur l'algorithme de la retro propagation pour rétro propager l'erreur entre l'état des neurones de sortie et la réponse désirée sur les poids des connexions des couches antérieures.

6.3 L'algorithme d'apprentissage

L'algorithme d'apprentissage est le suivant [DMS⁺08] :

1. Initialisation des poids du réseau
2. Présentation du vecteur d'apprentissage à l'entrée du réseau (couche d'entrée)
3. Calcul du vecteur de sortie S : le système propage les activités neuronales (état excité ou inhibé des neurones) à travers le réseau. Chaque neurone calcule la somme pondérée de ses entrées et transmet le résultat par une fonction de type sigmoïde pour produire sa valeur de sortie. Le vecteur de sortie S est le résultat du calcul de la dernière couche (couche de sortie).
4. Calcul de l'erreur. On compare alors les valeurs de sortie actuelles S avec les valeurs désirées de R appelé vecteur de référence (ou encore vecteur cible).
On en déduit l'erreur J , définie comme la somme des carrés des différences observée sur chaque unité de la couche de sortie.
5. Rétro propagation de l'erreur : l'algorithme de rétro propagation du gradient permet récursivement de rétro propager l'erreur de la couche de sortie vers les couches cachées jusqu'à la première couche du réseau.
6. Modification des poids : les poids de chaque neurone sont modifiés, soit à chaque présentation d'un vecteur d'apprentissage (méthode du gradient stochastique), soit après cumul de l'erreur pour un certain nombre de vecteurs d'apprentissage (méthode du gradient standard).

6.4 Les avantages et les inconvénients des RNs

LES AVANTAGES	LES INCONVÉNIENTS
Classifieur très précis (si bien paramétré).	Détermination de l'architecture des réseaux est complexe.
Apprentissage automatique des poids.	Paramètres difficiles à interpréter (boite noire).
Possibilité de faire parallélisme (les éléments de chaque couche peuvent se fonctionné en parallèle).	Difficulté de paramétrage surtout pour le nombre de neurones dans la couche cachée.
Résistance aux pannes (si un neurone ne fonctionne plus, le réseau ne se perturbe pas).	

TABLE 3 – Avantages et inconvénients des RNs

7 Application

7.1 Classification sans réduction

Cette étape consiste à faire une classification avec l'échantillon complet c'est-à-dire toute la base de données, alors nous avons effectué une répartition de l'échantillon entre deux ensembles : 2/3 pour l'ensemble d'apprentissage et 1/3 pour l'ensemble du test. Pour l'évaluation d'un classifieur on juge dans la plus part des cas les critères classiques : le taux de classification, le taux d'erreur, la sensibilité et spécificité.

Après avoir expérimenté plusieurs testes on a trouvé que la meilleure architecture du réseau qui donne un résultat important que les autres est de : Prendre une seule couche cachée avec 8 neurones et une fonction d'activation par défaut donc les résultats se résume dans la Table 4 :

Caractéristiques	22
Nombre de neurones dans la couche d'entrée	22
Nombre de couche cachée	1
Nombre de neurones dans la couche cachée	8
Taux de classification%	90.62
Sensibilité%	62.5
Spécificité%	100

TABLE 4 – Performances de classification sans réduction

7.2 Classification avec réduction (ACI)

Étape de réduction

A l'aide de l'algorithme FastICA [Alg10] sur MATLAB on a pus extraire les paramètres les plus

pertinentes via une interface qui facilite la réduction. On a réalisé une réduction : pour 3, 5, 10 et 15 Composantes indépendantes, et dans chacune on a pu avoir les différentes variables pertinentes.

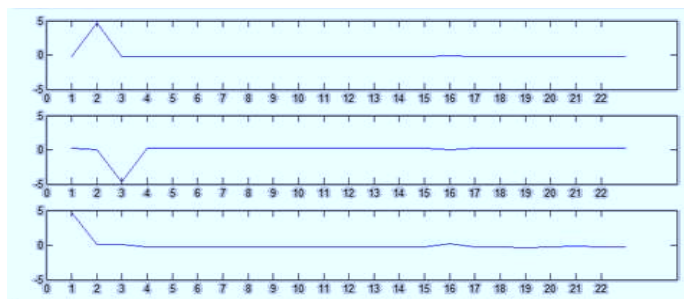


FIGURE 5 – Réduction à 3 Composantes indépendantes

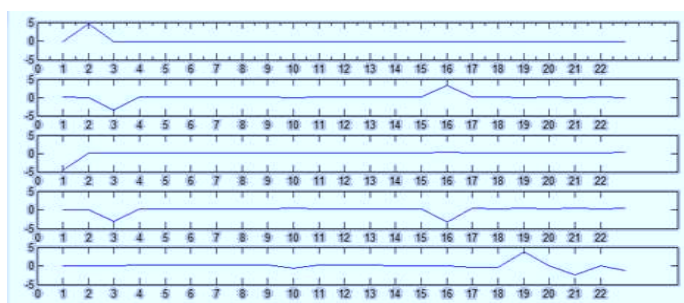


FIGURE 6 – Réduction à 5 Composantes indépendantes

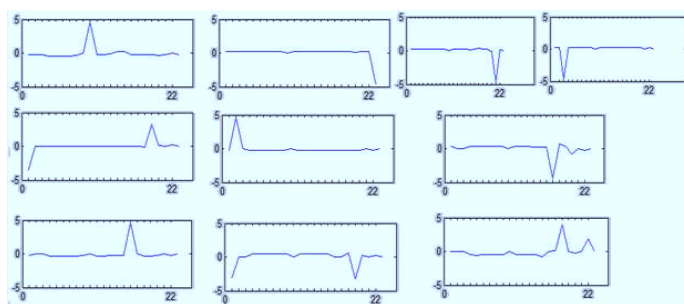


FIGURE 7 – Réduction à 10 Composantes indépendantes

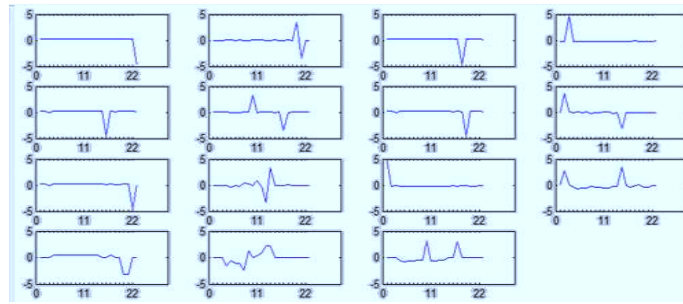


FIGURE 8 – Réduction à 15 Composantes indépendantes

Étape de Classification

L'étape consiste à faire aussi une classification mais avec les résultats obtenus de la réduction de données c'est-à-dire les paramètres pertinents extraient de l'analyse en composantes indépendantes qui représentent les neurones de la couche d'entrée du réseau de neurones. Il a la même architecture que le précédent pour que la comparaison de résultats soit logique aussi on a réalisé la phase du test avec la même matrice.

Donc on a effectué une classification pour les quatre échantillons juste pour confirmer nos résultats d'application se résume dans la Table 5 :

Caractéristiques	Les composantes indépendantes			
	3 CI	5 CI	10 CI	15CI
ECHANTILLONS	3	8	9	16
NOMBRE DE PARAMÈTRES	1	1	1	1
NOMBRE DE COUCHE CACHÉE	8	8	8	8
NOMBRE DE NEURONES DANS LA COUCHE CACHÉE	93.75	92.81	85.93	81.25
TAUX DE CLASSIFICATION%	75	68	56.25	75
SENSIBILITÉ%	100	100	100	100
SPÉCIFICITÉ%				

TABLE 5 – Performances de classification avec réduction

8 Discussion des résultats

La classification avec réseau de neurone sans réduction n'a pas donné des résultats très intéressants mais elle nous a montré que les premiers échantillons sont beaucoup plus efficaces pour le test que pour l'apprentissage.

L'étape de réduction de données nous montre clairement que chaque échantillon a donné des paramètres pertinents mais l'idée étant d'aller plus dans l'interprétabilité de notre application est de faire une extraction de ces composantes indépendantes des variables les plus pertinentes les englobants suivant les figures (Figure 5-6-7-8). Une manière de confirmation puisque le choix est très important pour la suite du travail. Les résultats obtenus sont les suivants :

NOMBRE DE COMPOSANTES	NOMBRE DE PARAMÈTRES	LES PARAMÈTRES
3	3	MDVP :Fo(Hz) MDVP :Fhi(Hz) MDVP :Flo(Hz)
5	8	MDVP :Fo(Hz),MDVP :Fhi(Hz),MDVP :Flo(Hz), MDVP :Shimmer(dB), HNR, spread1, D2,PPE
10	9	MDVP :Fo(Hz),MDVP :Fhi(Hz),MDVP :Flo(Hz), MDVP :Shimmer(dB), HNR, RPDE,DFA, spread1, D2.
15	16	MDVP :Fo(Hz),MDVP :Fhi(Hz),MDVP :Flo(Hz), Jitter :DDP,MDVP :Shimmer, MDVP :Shimmer(dB),MDVP :APQ,Shimmer :DDA,NHR, HNR, RPDE,DFA, spread1, spread2, , D2,PPE.

TABLE 6 – Tableau des paramètres pertinents

La classification avec la réduction de données par ACI a donné de meilleurs résultats, un taux de 93.75 % ; avec les trois composantes indépendantes où nous avons pu extraire les variables pertinentes qui tournent sur la fréquence fondamentale de la voix, cette dernière a été aussi efficace pour des études similaires par exemple si une personne fume ou pas , aussi on peut dire que le choix de ACI était judicieux pour une telle base de données puisque nous avons eu une amélioration en taux de classification par rapport à ce qui a été fait le mathématicien Max Little [TLM⁺12],[TLMR10] qui a arrivé à un taux de 86% par la classification avec la méthode du vecteur à support de marges SVM.

Caractéristiques	Sans ACI	Avec ACI
Nombre de neurones dans la couche d'entrée	22	3
Nombre de couche cachée	1	1
Nombre de neurones dans la couche cachée	8	8
Taux de classification%	90.62	93.75
Sensibilité%	62.5	75
Spécificité%	100	100

TABLE 7 – Performances de classification avec et sans réduction

Et si on veut comparer entre la classification avec réduction et la classification sans faire une réduction, il est préférable de la faire pour simplifier l'étude et surtout pour obtenir plus de robustesse.

9 Conclusion

Un classifieur neuronal peut être un outil pratique dans le processus de diagnostic. Il se compose de règles qui sont faciles à interpréter par l'expert humain. Le réseau de neurone est un modèle dont la conception est inspirée du fonctionnement des neurones biologiques. Ceci est très important pour les systèmes de prise de décision, car les experts n'acceptent pas une évaluation sur or-

dinateur, à moins qu'ils comprennent pourquoi et comment une recommandation a été donnée.

Notre objectif principal était d'améliorer le taux de classification par rapport au préalable jusqu'à 93% donc le réseau de neurone a fait ressortir aussi son efficacité dans sa capacité de classification des données non supervisées pour un apprentissage plus rapide avec une structure réduite, par le biais des méthodes d'optimisation ou de réductions des paramètres tel que l'analyse en composantes indépendantes.

Ce rapport de recherche constitue un premier pas dans le développement d'un système d'aide au diagnostic médical intelligent qui peut aider les experts dans le domaine d'une détection précoce de parkinson.

Références

- [Alg10] The FastICA Algorithm. The fastica package for matlab. 2010.
- [Car02] Jean-François Cardoso. Analyse en composantes indépendantes. In *Actes des XXXIV^{èmes} Journées de Statistique, JSBL 2002, Bruxelles*, 2002.
- [Dep12] Cédric Depond. La maladie de parkinson diagnostiquée à la voix. *Techno-science*, <http://www.techno-science.net/?onglet=news&news=10555> :Access 05/03/2013, 01/07/2012.
- [DMS⁺08] Gérard Dreyfus, Jean-Marc Martinez, Mannuel Samuelides, Fouad Badran Mirta Gordon, and Sylvie Thiria. *Apprentissage statistique : réseaux de neurones, cartes topologiques, machines à vecteurs supports*. 2008.
- [HO00] A. Hyvarinen and E. Oja. Independent component analysis : algorithms and applications. *Neural Netw*, 13 :411–430, 2000.
- [Lit08] Max Little. UCI machine learning repository. 2008.
- [TdAL09] Fabien Tschirhart and (dir. Alain Lioret). Réseaux de neurones formels appliqués à l'intelligence artificielle et au jeu. Master's thesis, ESGI (mémoire de master de recherche en multimédia et animation numérique), Paris, 2009.
- [TLM⁺12] A. Tsanas, M.A. Little, P.E. McSharry, J. Spielman, and L.O. Ramig. Novel speech signal processing algorithms for high-accuracy classification of parkinson's disease. *IEEE Transactions on Biomedical Engineering*, 59(5) :1264–1271, 2012.
- [TLMR10] A Tsanas, M.A. Little, P.E. McSharry, and L.O. Ramig. Nonlinear speech analysis algorithms mapped to a standard metric achieve clinically useful quantification of average parkinson's disease symptom severity. *Journal of the Royal Society Interface*, 8(59) :842–855, 2010.
- [Vas40] Flore Vasseur. Diagnostiquer parkinson, bientôt simple comme un coup de fil. *LE MONDE*, http://www.lemonde.fr/a-la-une/article/2012/08/08/diagnostiquer-parkinson-bientot-simple-comme-un-coup-de-fil_1743267_3208.html :Access 15/03/2013, 08.08.2012 à 16h40.