

Université du Québec en Outaouais
Département d'informatique et d'ingénierie

**Diagnostic de la dépression et prédiction de l'efficacité d'un
traitement chez des patients atteints d'un désordre dépressif majeur**

Mémoire
présenté comme exigence partielle à la maîtrise en sciences et technologies de
l'information

Par
Ilhem Rima

Directrice de recherche :

Dr. Rokia Missaoui, UQO

Membres du Jury :

Dr Ana-Maria Cretu, UQO

Dr Morad Benyoucef, École Telfer de gestion, université d'Ottawa

Avril 2014

Remerciements

Je voudrais remercier ma directrice de recherche Dr. Rokia missaoui pour m'avoir guidée, conseillée, et encouragée tout au long de ma thèse. Mes remerciements vont également à Dr. Verner Knott, directeur du Laboratoire de Neuroélectrophysiologie et de Recherches Cognitives (LNRC) de l'Institut de recherche en santé mentale de l'université d'Ottawa, pour la documentation, les explications et l'aide financière fournies.

Table des matières

Remerciements	I
Résumé	iii
1 Introduction	1
2 État de l'art	5
2.1 Notions fondamentales de la fouille de données	5
2.2 Tâches de fouille de données.....	6
2.2.1 Classification	6
2.2.2 Regroupement	7
2.2.3 Association	7
2.3 L'électroencéphalogramme	8
2.4 Marqueurs biologiques électroencéphalographiques.....	9
2.5 Modèles de fouille de données.....	13
2.5.1 Réseaux de neurones.....	13
2.5.2 Regroupement	15
2.5.3 Règles d'association.....	16
2.5.4 Arbre de décision	17
3 Méthodologie	20

3.1 Prétraitement des données	20
3.2 Démarche de fouille de données.....	21
4 Fouille de données – Partie 1	24
4.1 Prédiction	24
4.1.1 Réseaux de neurones	24
4.1.2 Règles d’association	27
4.1.3 Regroupement	29
4.2 Diagnostique de la dépression.....	30
4.2.1 Règles de classification.....	30
4.2.2 Réseaux de neurones	32
5 Fouille de données – Partie 2	34
5.1 Classification par arbre de décision	34
5.2 Règles d’association avec négation et généralisation de variables.....	48
6 Conclusion	53
Références	55

Résumé

Le but de cette recherche est d'utiliser des techniques de fouille de données sur des informations extraites d'électroencéphalogrammes (EEG) d'individus pour aussi bien le diagnostic de la dépression que la prédiction de l'efficacité d'un traitement antidépresseur à une étape précoce des soins. Pour le volet diagnostique, il s'agit de comparer deux groupes d'individus (malades dépressifs et sujets sains) sur la base d'ondes cérébrales (alpha, bêta, thêta et delta) extraites des EEG pour identifier celles qui expliquent le mieux l'état dépressif et pour générer des règles de classification permettant de prédire la classe d'appartenance d'un individu (sain ou dépressif). Quant au volet de prédiction, il s'agit d'utiliser des ondes cérébrales et éventuellement d'autres mesures cliniques afin d'identifier les patients qui ne répondront pas favorablement à un traitement après juste une semaine de soins au lieu de douze semaines, comme il est souvent d'usage. Cela éviterait de faire subir aux patients non répondeurs les effets indésirables d'un traitement inefficace sur une longue durée.

Après quelques transformations de données EEG, nous avons appliqué trois techniques de fouille de données : la classification, le regroupement et les règles d'association, et avons obtenu des résultats jugés intéressants et utiles par nos partenaires cliniciens. Certains résultats confirment ceux déjà obtenus par d'autres recherches antérieures et portent principalement sur l'impact des ondes thêta sur la prédiction précoce de l'efficacité d'un traitement antidépresseur. Nous avons également adapté des procédures de fouille de données développées au laboratoire LARIM de l'UQO aux données sur la dépression afin de générer des motifs spéciaux comme les règles d'association avec négation ou les règles généralisées impliquant la prise en compte de hiérarchies de généralisation sur les attributs.

Ce travail rentre dans le cadre d'un projet de recherche conduit au sein du CBAP (*IBM centre for Business Analytics and Performance*) de l'École Telfer de gestion en collaboration avec Docteur Verner Knott, directeur du Laboratoire de Neuroélectrophysiologie et de Recherches Cognitives (LNRC) de l'Institut de recherche en santé mentale de l'université d'Ottawa.

Chapitre 1 Introduction

Les progrès technologiques rapides dans la collecte et le stockage de données ont permis aux organisations d'accumuler de grandes quantités de données dans différents domaines. Comme les données recueillies ainsi que leurs formes diverses (données spatio-temporelles, géographiques, bases de données multimédia, etc.) peuvent être très volumineuses et complexes, les outils d'analyse de données classiques peuvent s'avérer inefficaces dans l'exploration de ces données.

Avec l'avènement de la fouille de données, les techniques d'analyse de données traditionnelles qui se fondent en grande partie sur des méthodes statistiques ont été combinées avec des algorithmes sophistiqués pour un traitement efficace de grands volumes de données ainsi qu'une aide à l'interprétation et une visualisation conviviale des résultats.

La fouille de données est le processus d'extraction de connaissances inconnues à partir de bases ou d'entrepôts de données. Le résultat d'une fouille de données représente de la connaissance qui peut prendre la forme d'associations (ex. si un client achète le produit A, alors il achète les articles B et C), de groupes formés par des techniques de regroupement (*clustering*), de classes ou d'une manière générale de motifs (*patterns*).

La fouille de données constitue aujourd'hui une partie intégrale d'un processus de veille économique (*business intelligence*) dans différents domaines d'application comme le domaine médical, industriel, ou commercial. Par exemple, dans le domaine commercial, nous sommes confrontés tous les jours à de la publicité ciblée vers les clients qui sont les plus susceptibles de répondre à une sollicitation. Les détaillants appliquent aussi à leur tour des techniques de fouille aux données recueillies en temps réel sur les achats de leurs clients pour mieux comprendre leurs besoins et faire des suggestions d'achat d'articles. Par ailleurs, la fouille de données joue un rôle important dans le domaine scientifique. C'est le cas par exemple des données spatio-temporelles recueillies par les satellites de la NASA qui ne peuvent être analysées avec les techniques traditionnelles à cause de la taille, la nature et la vélocité des données. En utilisant la fouille de données, les scientifiques ont la possibilité de découvrir de nouveaux motifs inattendus ou confirmer/infirmier des hypothèses préalablement énoncées. La fouille de données a aussi réussi à faire ses preuves dans le domaine médical avec ses différentes spécialités telles que la psychiatrie.

Ce projet de recherche consiste en l'utilisation de la fouille de données comme démarche d'analyse de données afin d'extraire des connaissances, auparavant inconnues par les cliniciens, à partir des électroencéphalogrammes (EEG) de patients dans le but ultime d'une part, de poser

un diagnostic de dépression et d'autre part, de prédire l'efficacité d'un traitement antidépresseur à une étape précoce du soin des patients.

Une réponse partielle ou inadéquate à un traitement antidépresseur est commune dans les troubles dépressifs majeurs (*major depressive disorders*). Seulement près de 30% à 40% des patients qui reçoivent une pharmacothérapie appropriée atteignent une rémission complète (absence ou quasi-absence de symptômes). C'est la raison pour laquelle les cliniciens se trouvent contraints de faire plusieurs tentatives et prescrire différents médicaments antidépresseurs avant d'atteindre une guérison complète de leurs patients [1].

Par ailleurs, les patients sont appelés à subir les effets secondaires de chaque nouvel antidépresseur prescrit, ce qui engendre un véritable fardeau pour les patients sur le plan socio-économique et peut même mener vers le suicide.

Jusqu'à présent, les cliniciens ont du mal à prédire la réponse d'un patient à un traitement donné avant ou peu après son application. À l'heure actuelle, les cliniciens se basent principalement sur une évaluation subjective des symptômes cliniques avant l'instauration du traitement (*base line*) ainsi qu'une semaine après le début du traitement.

Le but de notre travail est donc de déterminer pour les cliniciens des marqueurs biologiques (indicateurs d'un état pathologique donné extraits d'électroencéphalogrammes) qui serviront pour le *diagnostic* de la dépression et la *prédiction* de l'efficacité d'un traitement. Ainsi, il serait possible aux cliniciens d'identifier les patients qui ne répondent pas favorablement à un traitement au plus tard une semaine après son démarrage. Cela permet d'éviter de faire subir à ces patients les effets indésirables d'un traitement inefficace et donc les faire souffrir pendant une période de douze semaines sans qu'il soit possible d'obtenir une rémission complète.

Contrairement aux variables cliniques utilisées jusqu'à présent par les médecins pour l'évaluation du traitement par antidépresseurs lesquels ne sont pas toujours fiables, les nombreuses recherches effectuées à ce sujet indiquent l'existence de preuves croissantes que les mesures basées sur les réactions électriques cérébrales aux médicaments peuvent mieux discriminer entre les patients répondeurs et ceux non répondeurs à un traitement donné, et ceci à un stade précoce des soins. En plus de l'utilisation potentielle de l'EEG comme *prédicteur* objectif précoce de la réponse aux antidépresseurs, l'EEG est aussi en mesure de fournir des données directement liées à l'activité neuronale qui peut être anormale dans certains troubles et donc de servir comme outil de *diagnostic* objectif pour différencier entre les patients dépressifs et les sujets sains.

L'électro-encéphalographie est une méthode d'exploration cérébrale qui mesure l'activité électrique du cerveau par des électrodes placées sur le cuir chevelu. Une fois enregistrée, cette activité se présente sous forme d'un tracé appelé *électroencéphalogramme*.

L'activité EEG à chaque électrode du cuir chevelu reflète la sommation de l'activité électrique synchrone de milliers ou de millions de neurones localisés sur le cortex cérébral dans les zones entourant l'électrode.

Notre recherche comporte donc deux parties : une partie diagnostic et une partie prédiction avec un accent plus prononcé sur cette dernière. La partie diagnostic consiste à définir les caractéristiques spécifiques d'un profil dépressif d'un cerveau donné en comparant l'EEG des patients dépressifs avec celui des sujets sains. La partie prédiction consiste à déterminer si les valeurs de l'EEG recueillies avant et pendant la première semaine du traitement peuvent être utilisées pour prédire la réponse au traitement après douze semaines.

Ce travail est mené dans le cadre des activités de recherche du centre CBAP (*IBM centre for Business Analytics and Performance*) qui est un laboratoire de recherche en analyse d'affaires (*Business Analytics*), en collaboration avec Docteur Verner Knott, directeur du Laboratoire de Neuroélectrophysiologie et de Recherches Cognitives (LNRC) de l'Institut de recherche en santé mentale de l'université d'Ottawa.

Le laboratoire de recherche CBAP se compose d'un groupe multidisciplinaire de chercheurs d'IBM, d'organismes fédéraux, de l'École de gestion Telfer et de l'École de science informatique et de génie électrique de l'université d'Ottawa. Ce travail implique également ma directrice de recherche Professeure Rokia Missaoui ainsi que son université d'attache.

Les données utilisées dans le cadre de cette recherche nous ont été fournies par le LNRC. Il s'agit de l'EEG absolue de 52 patients qui ont subi un traitement de six semaines par deux antidépresseurs, à savoir l'*escitalopram* et le *bupropionet*, et de 43 sujets sains. Les données d'une autre mesure EEG appelée *cordance* nous ont aussi été fournies mais à une date plus récente. Ces données concernent 39 patients qui ont subi le même traitement et l'EEG des patients a été enregistré avant le début et une semaine après le traitement.

À ce jour, on peut dire que les résultats des différentes études utilisant l'EEG pour prédire la réponse au traitement du désordre dépressif majeur sont insuffisants et ne peuvent être généralisés pour être utilisés comme prédicteurs. C'est la raison pour laquelle on pense que notre travail va apporter une contribution à l'interprétation de ces résultats en utilisant des techniques de fouille de données tout en nous basant sur les suggestions de variables à privilégier et sur des hypothèses émises par l'équipe de Dr Knott.

Le présent mémoire est organisé comme suit. Dans la section 2, nous présentons le contexte du projet en introduisant les notions fondamentales sur la fouille de données, la technologie de l'électroencéphalographie et un survol de la littérature sur l'utilisation potentielle de l'EEG comme outil de diagnostic et de prédiction. La section 3 décrit la démarche méthodologique alors que la section 4 fournit une présentation et une interprétation des premiers résultats obtenus.

La section 5 décrit les récents résultats obtenus en utilisant principalement les données sur les ondes thêta et en misant sur l'introduction de la généralisation et la négation dans le processus de la fouille de données. Finalement, la section 6 conclut ce mémoire.

Chapitre 2

État de l'art

Dans ce chapitre, nous allons tout d'abord faire un rappel sur les notions fondamentales de la fouille de données pour ensuite faire un survol de la littérature sur l'utilisation potentielle de l'électroencéphalogramme comme à la fois un prédicteur objectif précoce de la réponse aux antidépresseurs et un outil de diagnostic.

2.1 Notions fondamentales de la fouille de données

Comme nous l'avons déjà mentionné en introduction, la fouille de données [2] est le processus d'extraction de connaissances inconnues à partir de bases ou d'entrepôts de données. Le résultat de la fouille de données peut prendre la forme d'associations, de concepts (groupes), de classes ou d'une manière générale de motifs (*patterns*) en utilisant des techniques de classification, de regroupement, de fouille de règles d'association et même de techniques de détection de valeurs aberrantes ou exceptionnelles (*outliers*).

Même si les données stockées dans les bases ou entrepôts de données de type numérique ou qualitatif sont les plus souvent utilisées pour la fouille de données, cette dernière a été étendue à des types non conventionnels de données telles les données graphiques, temporelles, spatiales, textuelles, et multimédia. Par exemple, dans les données boursières, la fouille de données peut être appliquée à la détection des périodes les plus convenables pour l'achat d'une action. Un autre exemple est l'application de la fouille de données dans une vidéo de hockey pour détecter les séquences vidéo d'un but dans un match donné. En outre, la fouille de données peut être appliquée sur des données spatiales comme les cartes géographiques pour la détection de changements géographiques de la terre.

La fouille de données est une partie intégrale du processus de la découverte de connaissances à partir des bases de données (*knowledge discovery in databases*) qui a pour but d'extraire des connaissances à partir de données brutes. Ce processus se compose de trois principales étapes : la première étape est représentée par le prétraitement des données et vise à transformer les données d'entrée brutes en un format approprié pour la fouille de données. Le nettoyage et l'intégration de données de diverses sources fait aussi partie du prétraitement des données. De plus, cette étape inclut la sélection d'observations et de leurs propriétés pertinentes au projet de fouille de données en question. La deuxième étape consiste à appliquer des algorithmes de fouille de données. Le choix des algorithmes est fonction de l'objectif de la fouille de données. La troisième et dernière étape concerne le post-traitement et consiste à garantir que seuls les résultats valables et utiles sont présentés d'une manière conviviale à l'utilisateur final.

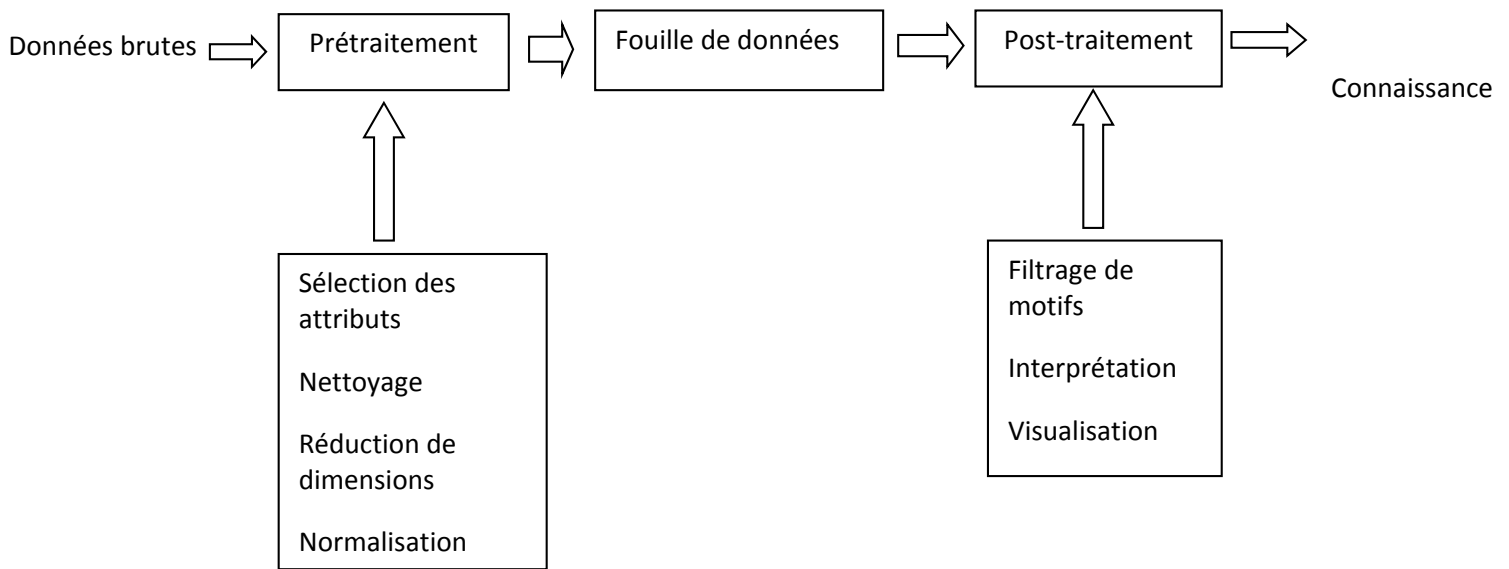


FIG. 1- Processus de découverte de connaissances à partir de bases de données.

2.2 Tâches de fouille de données

Il existe deux catégories de tâches de fouille de données : la prédiction et la description. Les tâches descriptives ont pour but d'identifier des motifs (corrélations, associations, tendances, groupes, trajectoires ou anomalies) cachés dans les données. Les tâches prédictives ont, quant à elles, pour but d'affecter un objet à une classe prédéfinie d'instances (classification) ou de prédire la valeur d'un attribut particulier sur la base de la valeur d'autres attributs (régression).

Les tâches de fouille de données sont très variées et distinctes et peuvent aboutir à divers types de motifs. L'utilisation d'une certaine tâche est donc basée sur le type de motifs que nous désirons trouver.

Parmi les tâches les plus répandues, on peut citer celles que nous avons retenues pour notre analyse : la classification, le regroupement (*clustering*) et les règles d'association.

2.2.1 Classification

La classification est une tâche prédictive qui relève de l'apprentissage dit *supervisé*. Alors qu'elle permet de prédire l'appartenance d'une observation (objet, individu, événement) à une classe

prédéfinie d'instances, la régression est aussi une tâche prédictive mais qui prédit la valeur d'un attribut en fonction d'autres attributs.

Un modèle de classification est construit à partir d'un ensemble d'apprentissage (*training set*) dans lequel chaque observation (un objet désigné par une ligne dans une table relationnelle d'une base de données) est représentée par un vecteur de ses attributs parmi lesquels un seul attribut désigne une classe d'appartenance (ex. sujet sain versus malade). Le modèle est construit en analysant la relation entre les attributs et la classe d'appartenance des observations. Une fois le modèle créé à partir de l'ensemble d'apprentissage, il est utilisé pour prédire la classe d'appartenance d'un objet. Par exemple, en utilisant un ensemble d'apprentissage contenant des données sur des sujets sains et des sujets malades, on peut construire un modèle qui peut conclure la maladie chez un patient à partir de données personnelles et médicales (âge, poids, pression artérielle, etc.).

Un modèle de classification peut prendre plusieurs formes telles qu'un arbre de décision, des règles de classification (règles de type *si ... alors* dérivées à partir de l'arbre de décision) ou un réseau de neurones [3].

2.2.2 Regroupement

Contrairement aux tâches prédictives, les tâches descriptives ont une nature exploratoire. Elles visent à identifier des motifs (corrélations, tendances, grappes ou anomalies) cachés dans les données.

Le regroupement (*clustering*) est une tâche descriptive qui permet de dégager des groupes homogènes d'objets en analysant les similarités des valeurs de leurs attributs. Contrairement à la classification, le regroupement analyse un ensemble de données sans spécification de la classe d'appartenance des objets, d'où la dénomination « apprentissage non supervisé ». Parmi les algorithmes de regroupement, c.-à-d. de formation de grappes/groupes (*clusters*), on cite la classification hiérarchique ascendante (ou descendante), la méthode de *k-means*, le regroupement conceptuel de Michalski et l'analyse formelle de concepts [4].

2.2.3 Association

L'association est une tâche descriptive qui relève de l'apprentissage non supervisé. Elle permet l'extraction de règles d'association par la découverte de corrélations, liens ou structures causales entre les ensembles d'attributs (ou items). Une règle d'association prend la forme suivante : *prémisse* –> *conclusion*. La fouille des règles d'association est utilisée dans divers domaines et applications du monde réel. Elle a été initialement appliquée dans l'analyse du panier du

consommateur pour identifier les articles qui sont fréquemment achetés ensemble en vue de mieux comprendre les besoins et les habitudes de magasinage des clients. Par exemple, on peut découvrir la règle *pain* – > *lait* qui signifie que les clients qui achètent du pain ont aussi tendance à acheter du lait [3, 5].

Pour notre analyse des données de la dépression, nous allons utiliser deux techniques de fouille prédictive de type classification : les réseaux de neurones et les arbres de décision, et deux méthodes de fouille descriptive : le regroupement par la méthode *k-means* et l'identification des règles d'association.

2.3 L'électroencéphalogramme

Afin de pouvoir comprendre les résultats de cette étude, il est important de définir quelques notions concernant l'électroencéphalogramme.

Nous rappelons que l'électroencéphalographie est une méthode d'exploration cérébrale qui évalue l'activité neurophysiologique du cerveau en mesurant les potentiels électriques au niveau des neurones par le biais d'électrodes placées sur le cuir chevelu. Cette activité neurologique se présente sous forme d'électroencéphalogramme.

L'activité électrique cérébrale rythmique qu'enregistre l'EEG est divisée en bandes de fréquences différentes. La bande alpha est divisée en trois fréquences : alpha total (8 à 14 Hz), alpha1 (8 à 11 Hz) et alpha2 (11 à 14 Hz). La bande alpha apparaît avec la fermeture des yeux et la relaxation, et s'atténue avec l'ouverture des yeux ou l'effort mental. La bande bêta a une fréquence comprise entre 12 et 30 Hz et apparaît lors de la réflexion ou d'une forte concentration. La bande thêta, avec une fréquence comprise entre 4 et 7 Hz, apparaît durant la somnolence ou la méditation. Finalement, la bande delta avec une fréquence atteignant jusqu'à 4 Hz apparaît chez les adultes sains pendant le sommeil profond.

La distribution des bandes de fréquence au niveau du cerveau varie avec l'âge, le niveau de vigilance, et est influencée par les médicaments et la pathologie cérébrale.

Il est à noter qu'il existe une asymétrie EEG entre l'hémisphère droit et l'hémisphère gauche du cerveau.

L'activité EEG est enregistrée à partir d'au moins seize électrodes placées sur le cuir chevelu (voir figure 1 pour l'emplacement des électrodes) avec les yeux fermés (EC) ou les yeux ouverts (EO). L'enregistrement des différents signaux est effectué par rapport à deux électrodes de référence appelées respectivement l'électrode mastoïde placée en dessous de l'oreille gauche et à droite sur l'os mastoïde et l'électrode Cz placée au centre du cuir chevelu.

Les signaux EEG sont également décrits en terme de puissance du signal électrique. L'EEG *absolu* désigne la puissance d'électricité mesurée dans une bande de fréquence EEG à une électrode donnée mesurée en microvolts carrés (μV^2). L'EEG *relatif* est le pourcentage de la puissance électrique contenue dans une bande de fréquence donnée par rapport à la puissance totale du spectre. Une autre mesure EEG appelée *cordance* est une mesure de l'activité cérébrale régionale qui combine l'EEG aussi bien absolu que relatif. Cette mesure présente une plus forte corrélation avec la perfusion cérébrale régionale par rapport à l'EEG absolu ou l'EEG relatif pris séparément [14]. Ainsi, l'intensité de cette cordance est fonction de l'activité régionale cérébrale mesurée par l'électrode.

Le temps de réalisation d'une séance d'enregistrement d'EEG dure trois à dix minutes. Une fois l'EEG enregistré, les résultats sont analysés par l'algorithme de transformée de Fourier rapide qui calcule la puissance des potentiels électriques des électrodes pour les différentes bandes de fréquences.

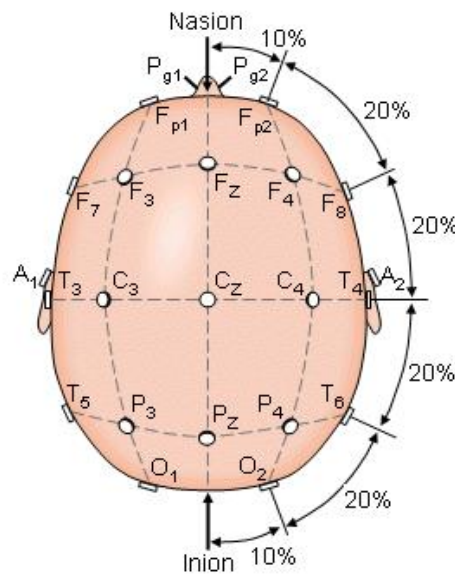


FIG. 2- Emplacement des électrodes sur le crâne.

Il faut signaler que l'électroencéphalographie présente plusieurs avantages par rapport à d'autres techniques qui sont utilisées dans ce domaine comme la neuro-imagerie. Il s'agit du faible coût et de la rapidité de production des EEG et des possibilités d'expansion à large échelle [1].

2.4 Marqueurs biologiques électroencéphalographiques

Comme mentionné en introduction, une réponse partielle ou inadéquate à un traitement antidépresseur est commune dans les désordres de dépression majeurs. À peine 30% à 40% des

patients qui reçoivent une pharmacothérapie (antidépresseurs) appropriée obtiennent une rémission complète (absence ou quasi-absence de symptômes).

Les groupes les plus communs d'antidépresseurs sont les antidépresseurs tricycliques (*Tricyclic Antidepressant*) et les inhibiteurs sélectifs de la recapture de la sérotonine (*Selective Serotonin Reuptake Inhibitors*).

Après de nombreuses recherches et expériences consécutives, plusieurs chercheurs croient de plus en plus que des méthodes basées sur l'EEG peuvent être utiles pour prédire la réponse des patients aux antidépresseurs et éventuellement pour guider les décisions thérapeutiques cliniques. Cependant, les chercheurs dans ce domaine reconnaissent qu'une telle démarche n'est pas encore très fiable du fait que c'est un domaine de recherche relativement nouveau et que les expériences réalisées sont encore à l'état préliminaire. Ils ont toutefois noté que les bandes offrant la meilleure distinction entre les patients répondeurs et non répondeurs aux antidépresseurs sont les bandes alpha et thêta.

Dans une étude menée par Knott et al. [6] sur cinquante patients ayant un désordre dépressif majeur et traités avec un antidépresseur SSRI (*Selective serotonin reuptake inhibitor*) appelé paroxétine pendant six semaines, il a été noté une augmentation de la force alpha avant le début du traitement et une diminution de la force thêta chez les patients répondeurs par rapport aux non répondeurs.

Dans une autre étude qui renforce les résultats précédents, Bruder et al. [7] ont rapporté que onze patients répondant au SSRI appelé fluoxétine démontraient une plus grande force alpha avant le début du traitement par rapport à sept patients non répondeurs et dix-huit sujets sains. En outre, ils ont observé que les répondeurs avaient une moindre activité alpha au niveau de l'hémisphère droit par rapport à l'hémisphère gauche du cerveau, contrairement aux patients non répondeurs qui montraient un résultat opposé.

Iosifu et al. [6] ont testé la capacité de la bande thêta à prédire la réponse au traitement. Dans une étude comprenant 82 patients atteints du syndrome de désordre dépressif majeur et traités avec des SSRI différents, la puissance de la bande thêta était plus basse chez les répondeurs après huit semaines de traitement.

En plus de l'utilisation de l'EEG absolu et relatif pour la prédiction de la réponse au traitement, des investigations ont également montré que la mesure de la *cordance* pourrait être utilisée comme prédicteur potentiel.

Des études antérieures faites sur des patients souffrant de dépression unipolaire ont montré que la diminution précoce de la cordance de l'EEG préfrontal au niveau de la bande thêta précédait l'amélioration des symptômes suite à la prise de médicaments antidépresseurs [8].

Dans une étude menée par Bares et al. [9], les chercheurs ont étudié l'impact de la diminution de

la cordance de l'EEG préfrontal (Fig. 3) après une semaine de traitement par la venlafaxine sur la prédiction de la réponse clinique de patients résistants.

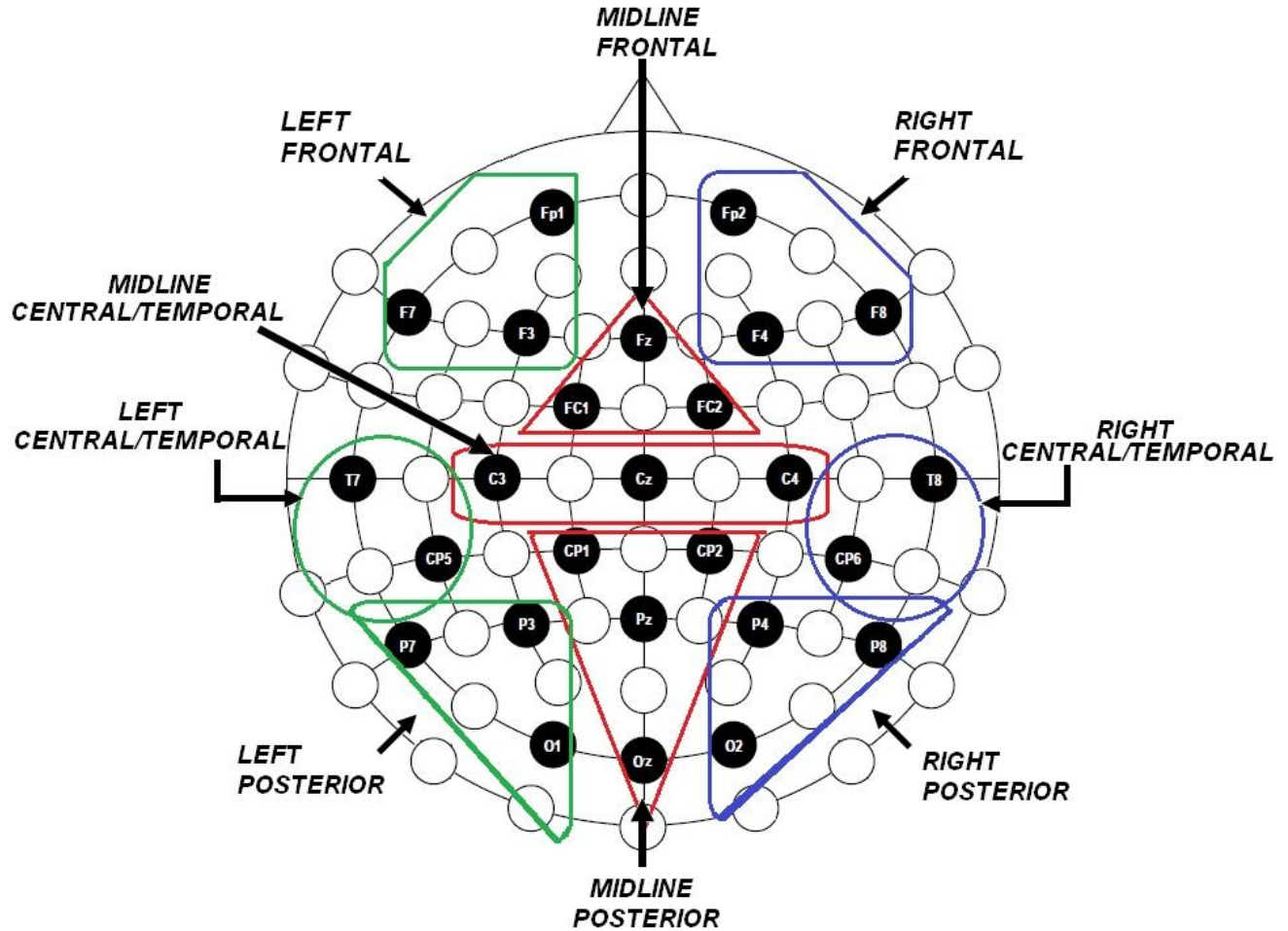


FIG. 3- Régions cérébrales

L'étude a été réalisée sur 25 patients hospitalisés qui ont terminé quatre semaines de traitement par la venlafaxine. La cordance a été enregistrée à la ligne de base (phase de prétraitement) et après une semaine de traitement. Les valeurs moyennes ont été calculées à partir des trois électrodes frontales (FP1, FP2 et Fz) pour chaque patient au niveau de la bande de fréquence thêta. Les résultats de l'étude ont montré que onze sur douze patients répondeurs et seulement cinq des treize patients non répondeurs ont montré une diminution de la cordance de l'EEG préfrontal après la première semaine de traitement.

La diminution de la cordance de l'EEG préfrontal après une semaine de traitement chez les patients répondeurs était importante. Cependant, il n'y avait pas de changement significatif pour les patients non répondeurs. Les valeurs prédictives de la réduction de la cordance de l'EEG préfrontal pour les patients répondeurs et non répondeurs ont été de 0.7 et 0.9 respectivement. Cette étude a montré que la réduction de la cordance de l'EEG préfrontal après une semaine de

traitement pourrait être un facteur prometteur pour prédire la réponse thérapeutique à la venlafaxine.

Les propriétés prédictives des régions MRF (*midline frontal and right frontal*) indiquées en figure 3 ont également été examinées. Cette partie du cerveau a été identifiée à partir de résultats physiologiques d'une enquête menée par Ian A. Cook et al. [10] sur des adultes qui n'ont jamais été déprimés ou exposés à des antidépresseurs, en mesurant la variation de la cordance au niveau de cette partie. Cette dernière a été la seule zone à présenter une diminution significative de la cordance après 48 heures puis une semaine de traitement.

Ces résultats ont amené ces chercheurs à examiner les propriétés prédictives de la cordance de la partie MRF chez des sujets déprimés et à déterminer une relation entre une diminution précoce de la cordance de la bande thêta et la rémission exprimée par un score final HAMD-D¹ ≤ 7. Ceci est dû au fait que les valeurs de la cordance de la bande thêta ont été systématiquement associées à la dépression ainsi qu'avec les effets thérapeutiques des médicaments antidépresseurs [11, 12, 13]. Une hypothèse suggère qu'une diminution précoce de la cordance de la bande thêta serait associée à la rémission des symptômes après traitement. L'étude a été réalisée sur 37 patients souffrant de désordres dépressifs majeurs dont 13 ont été traités à la fluoxétine (20 mg par jour) et 24 à la venlafaxine (150 mg par jour) pour une période de traitement de huit semaines. La moyenne des valeurs de la cordance aux électrodes FPZ, FZ, FP2, AF2, F4 et F8 a été calculée pour former la mesure de la cordance régionale MRF pour chaque patient.

Les changements dans la cordance de la bande thêta ont été enregistrés après 48 heures, une puis deux semaines après le début du traitement. La régression logistique a été utilisée pour évaluer ces changements. À la fin de la période de traitement, onze patients sur trente-sept avaient atteint une rémission complète. Il a été noté que la diminution de la cordance thêta au niveau de la région MRF après une semaine pouvait prédire une rémission avec 69% de précision. Ce résultat suggère qu'une rémission après instauration du traitement antidépresseur peut être prévisible en mesurant la cordance thêta de la région MRF après une semaine de traitement seulement.

Toutefois, des investigations supplémentaires sont nécessaires pour comparer les marqueurs biologiques et optimiser leur éventuelle utilisation dans la planification du traitement.

Dans une autre étude clinique menée par Rabinoff et al. [14], la cordance ainsi que l'EEG absolu et l'EEG relatif de chacune des 35 électrodes ont été employés. L'étude a été réalisée sur 51 adultes atteints de dépression unipolaire qui ont été traités avec la fluoxétine, la venlafaxine ou un placebo pendant une période de huit semaines. L'objectif de l'étude était d'évaluer l'utilité des arbres de décision de type CART pour faire une classification des patients répondeurs aux antidépresseurs et au placebo.

L'EEG a été enregistré pendant la phase prétraitement ainsi que 2, 7, 14, 28 et 56 jours après le

¹ HAMD-D: Hamilton Rating Scale for Depression.

début du traitement. Basé sur les principales constatations antérieures pour la région préfrontale, cette région a également été examinée par la moyenne des valeurs individuelles des trois électrodes préfrontales (FP1, FP2, FPZ) comme cela avait été fait dans les précédentes études [8,9].

Malgré le fait que la cordance et les valeurs absolues et relatives de l'EEG ont toutes été utilisées dans cette étude, seules la cordance de *la bande thêta* et la valeur absolue de l'EEG de *la bande alpha* ont été identifiées par l'arbre de décision comme étant significativement liées à la réponse des patients au traitement. En outre, parmi les 35 électrodes utilisées, seules quatre d'entre elles (FP1, AF2, PO2 et T4) ont été sélectionnées par l'arbre de décision. Notons que toutes ces électrodes à l'exception de FP1 provenaient de l'extérieur de la région préfrontale. Cependant, l'électrode FP1 a été révélée par l'arbre de décision comme étant l'élément le plus important et donc utilisée comme le critère de fractionnement primaire dans l'arbre pour la fluoxétine et la venlafaxine. Cela fait confirmer que l'utilisation des données physiologiques préfrontales est utile pour prédire le résultat clinique [14].

Les nombreuses études et expériences réalisées dans ce domaine prouvent que l'électroencéphalographie continue à être prometteuse pour prédire la réussite d'un traitement antidépresseur chez les patients présentant un désordre dépressif majeur.

Comme nous l'avons déjà mentionné, le but de notre travail est de fournir aux cliniciens des marqueurs biologiques qui serviront comme prédictifs objectifs précoces de la réponse des patients au traitement et cela implique deux volets : le diagnostic et la prédiction. L'aspect diagnostic consiste à comparer l'EEG des patients déprimés avec celui des sujets sains en vue de définir les caractéristiques spécifiques d'un profil dépressif d'un cerveau donné. L'aspect prédiction consiste à déterminer si les valeurs de l'EEG recueillies avant et après la première semaine de traitement peuvent être utilisées pour prédire la réponse à un traitement d'une durée de douze semaines.

2.5 Modèles de fouille de données

Tel qu'indiqué précédemment, la fouille de données va être effectuée en utilisant deux modèles prédictifs : les réseaux de neurones et les arbres de décision, et deux modèles descriptifs qui sont les règles d'association et le regroupement. Les modèles sont introduits dans la section ci-après.

2.5.1 Réseaux de neurones

Un réseau de neurones est un modèle mathématique prédictif qui peut être utilisé pour la classification ou la régression. Ce modèle est très schématiquement inspiré du fonctionnement

des neurones biologiques. Un réseau de neurones se compose de trois couches de nœuds : une couche d'entrée, une couche de sortie et une ou plusieurs couches cachées. Les nœuds de la couche d'entrée correspondent aux valeurs d'attributs d'entrée et les nœuds dans la couche de sortie représentent les valeurs de l'attribut prévisible pour le modèle. Entre la couche d'entrée et la couche de sortie se trouvent la/les couches cachées. Généralement un réseau de neurones peut avoir entre zéro et cinq couches cachées. Les nœuds dans la couche cachée reçoivent les valeurs des nœuds de la couche d'entrée avec leurs poids et fournissent des valeurs aux nœuds de la couche de sortie par le biais d'une fonction d'activation (*activation function*). Un poids décrit la pertinence ou l'importance d'une valeur d'entrée pour un nœud dans la couche cachée. Plus le poids associé à une valeur d'entrée est grand, plus cette valeur est importante pour faire une discrimination entre les attributs prévisibles. Chaque nœud du réseau est connecté aux nœuds de la couche adjacente de ce dernier comme le démontre la figure3.

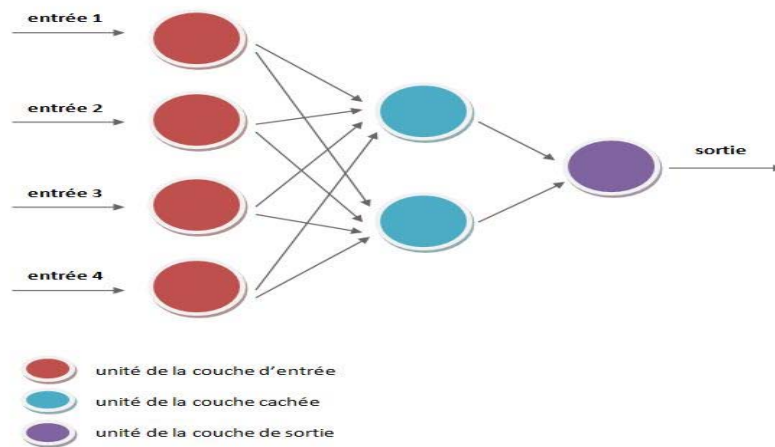


FIG. 4- Réseau de neurones.

Comme tous les modèles prédictifs relèvent de l'apprentissage supervisé, un ensemble d'apprentissage est utilisé pour généraliser la relation entre les attributs d'entrée et l'attribut prévisible. Une fois créé, le modèle peut être exploité pour faire des prédictions. Dans notre cas, l'ensemble d'apprentissage concerne des données extraites des EEG des patients et des sujets sains. Une fois les données fournies, l'objectif du modèle réseau de neurones est d'associer chaque état possible des attributs d'entrée à chaque état possible de l'attribut prévisible et de calculer la probabilité d'occurrence de l'attribut prévisible pour chaque état de valeur de l'attribut d'entrée.

Le module de services d'analyse (*Analysis Services*) du système MS SQL Server 2008 que nous utilisons en analyse de données fournit les informations suivantes au niveau des réseaux de neurones :

- **Score d'une variable** : importance de cette variable dans la prédiction de la classe d'appartenance.
- **Probabilité d'une valeur (ex. oui ou non)** : probabilité d'avoir ce résultat pour cette valeur.

Le tableau suivant contient un exemple des résultats générés par un réseau de neurones.

Attribut	Valeur	Favorise :non	Favorise :oui			
Salaire	120 000 – 140 000		<table border="1"> <tr> <td>Score: 100</td> </tr> <tr> <td>Probabilité du 'Non': 48.06 %</td> </tr> <tr> <td>Probabilité du 'Oui': 51.94 %</td> </tr> </table>	Score: 100	Probabilité du 'Non': 48.06 %	Probabilité du 'Oui': 51.94 %
Score: 100						
Probabilité du 'Non': 48.06 %						
Probabilité du 'Oui': 51.94 %						

Tableau 1- Résultats d'un réseau de neurones

Dans cet exemple de marketing ciblé où l'on cherche à déterminer quels sont les clients qui vont potentiellement acheter un produit, on peut voir que lorsque le salaire d'une personne est compris entre 120 000 et 140 000 dollars par année, il y a une probabilité de 48.06% qu'il n'achète pas le produit et une probabilité de 51.94% qu'il l'achète. Le score de 100% indique que le salaire a un impact décisif sur la classification.

2.5.2 Regroupement

Tel qu'indiqué auparavant, le regroupement est un modèle descriptif qui relève de l'apprentissage dit *non supervisé* et permet de dégager des groupes homogènes d'objets sur la base de leurs valeurs d'attributs. La distinction entre les différents regroupements est fonction de la similitude entre les objets. L'existence d'une grande similitude entre les objets au sein d'un groupe donné entraîne une grande différence entre les différents regroupements et donc une meilleure distinction entre ces groupes. Puisque ce modèle est de nature exploratoire, ces groupes ne sont pas prédéfinis auparavant mais plutôt découverts par l'algorithme de

regroupement. Le regroupement est une technique importante pour l'analyse et l'exploration des données pour un certain nombre de raisons dont la répartition des objets ou individus en groupes homogènes et la découverte de cas exceptionnels comme par exemple un patient qui a un profil totalement différent du reste des patients.

Il existe plusieurs méthodes de regroupement, chacune avec ses avantages et inconvénients. La méthode de regroupement utilisée dans notre recherche est la méthode *k-means*.

La méthode *k-means* est une méthode non hiérarchique de regroupement. Elle regroupe les données par utilisation d'une approche descendante qui commence par un nombre prédéfini de groupes et attribue au hasard une observation (objet) à chaque groupe (*cluster*). L'étape suivante consiste à comparer le reste des observations avec chacune des observations attribuées aux groupes en vue de les affecter au groupe auquel elles sont le plus similaires. Le point central (*central point*) de chaque groupe est ensuite calculé à l'aide d'une formule mathématique. Le processus de regroupement se poursuit par la détermination de la distance de chacune des observations à chaque point central des groupes. Si une observation est plus proche du centre d'un autre groupe, elle est déplacée vers ce groupe. Les centres des nouveaux et anciens groupes sont ensuite recalculés. Le processus de comparaison et de déplacement d'observations continue jusqu'à ce que le déplacement ne se justifie plus.

2.5.3 Règles d'association

Les règles d'association sont un autre exemple de modèle descriptif. L'objectif d'un algorithme de règles d'association est de découvrir des liens ou des associations entre les attributs des objets. Une règle d'association prend la forme suivante : $A \rightarrow B$ où A et B peuvent être des ensembles d'attributs décrivant des entités abstraites ou concrètes.

Un exemple d'une telle règle est comme suit :

Salaire > 120 000 \rightarrow achète voiture

Cependant, les règles d'association extraites à partir d'un ensemble de données peuvent ne pas être toutes utiles. Elles doivent être classées selon trois mesures qui sont le support, la confiance et l'intérêt (*lift*).

- La règle $r : A \rightarrow B$ a pour support le ratio S si S des objets de l'ensemble O possèdent les propriétés A et B .

$$S = \text{Prob}(A \cup B)$$

- La règle $r : A \rightarrow B$ a pour confiance C si la proportion C des objets de l'ensemble O ayant A ont également B .

$$C = \text{Prob}(B/A) = \text{Prob}(A \cup B) / \text{Prob}(B) = S / \text{Prob}(B)$$

Plus la valeur de C est proche de 1, plus la confiance est grande. Cependant, la confiance peut être trompeuse puisqu'elle n'indique pas la force de l'association entre la prémisse A et la conclusion B . Par exemple, dans un scénario d'analyse du panier du consommateur, si chaque transaction contient une bouteille d'eau, peut-être qu'elle est ajoutée à la charrette de chaque client automatiquement dans le cadre d'une promotion. Le modèle créerait dans ce cas plusieurs règles de confiance 1 prévoyant la présence de l'eau pour tout achat d'article. C'est la raison pour laquelle d'autres mesures de qualité ont été proposées.

- La règle $r : A \rightarrow B$ a pour intérêt I de la forme :

$$I = S / \text{Prob}(A) * \text{Prob}(B) = C / \text{Prob}(A)$$

Les valeurs d'intérêt supérieures à 1 indiquent une association positive tandis que les valeurs inférieures à 1 indiquent une association négative. Lorsque la valeur vaut 1, A et B sont indépendants et donc non corrélés.

2.5.4 Arbre de décision

Un arbre de décision est un autre exemple de modèle prédictif [5]. La classification à l'aide d'un arbre de décision permet d'affecter un objet à une classe prédéfinie d'instances en se basant sur un ensemble de valeurs d'attributs et sur un attribut de classification. Comme la classification relève de l'apprentissage dit supervisé, un arbre de décision est bâti selon un ensemble d'apprentissage. Cet ensemble est une table de valeurs discrètes comme le montre le tableau 1 où certains attributs sont des descripteurs et un attribut est l'attribut de classification. La figure 3 présente l'arbre de décision généré à partir de l'ensemble d'apprentissage du tableau 1 où chaque nœud non terminal de l'arbre spécifie un test d'attribut et chaque branche partant d'un nœud correspond à une valeur possible de l'attribut concerné. L'ordre de prise en compte des attributs se fait à l'aide d'une mesure statistique appelée "gain". Une fois le gain de chaque attribut est calculé, ces derniers sont placés dans l'arbre par ordre décroissant de gain. La racine de l'arbre a l'attribut avec le plus grand gain alors que les feuilles représentent la classification.

L'arbre de décision est un modèle utilisé dans plusieurs applications telles que l'approbation de crédit, le marketing ciblé, le diagnostic médical, l'analyse de l'efficacité d'un traitement, etc.

No	Revenu	Passé financier	Dettes	Risque
1	< 15K\$	mauvais	élevée	haut
2	15 à 35K\$	inconnu	élevée	haut
3	15 à 35K\$	inconnu	faible	moyen
4	< 15K\$	inconnu	faible	haut
5	> 35K\$	inconnu	faible	faible
6	> 35K\$	inconnu	faible	faible
7	< 15K\$	mauvais	faible	haut
8	> 35K\$	mauvais	faible	moyen
9	> 35K\$	bon	faible	faible
10	> 35K\$	bon	élevée	faible
11	< 15K\$	bon	élevée	haut
12	15 à 35K\$	bon	élevée	moyen
13	> 35K\$	bon	élevée	faible
14	15 à 35K\$	mauvais	élevée	haut

Tableau 2- Tableau de demandes de prêts

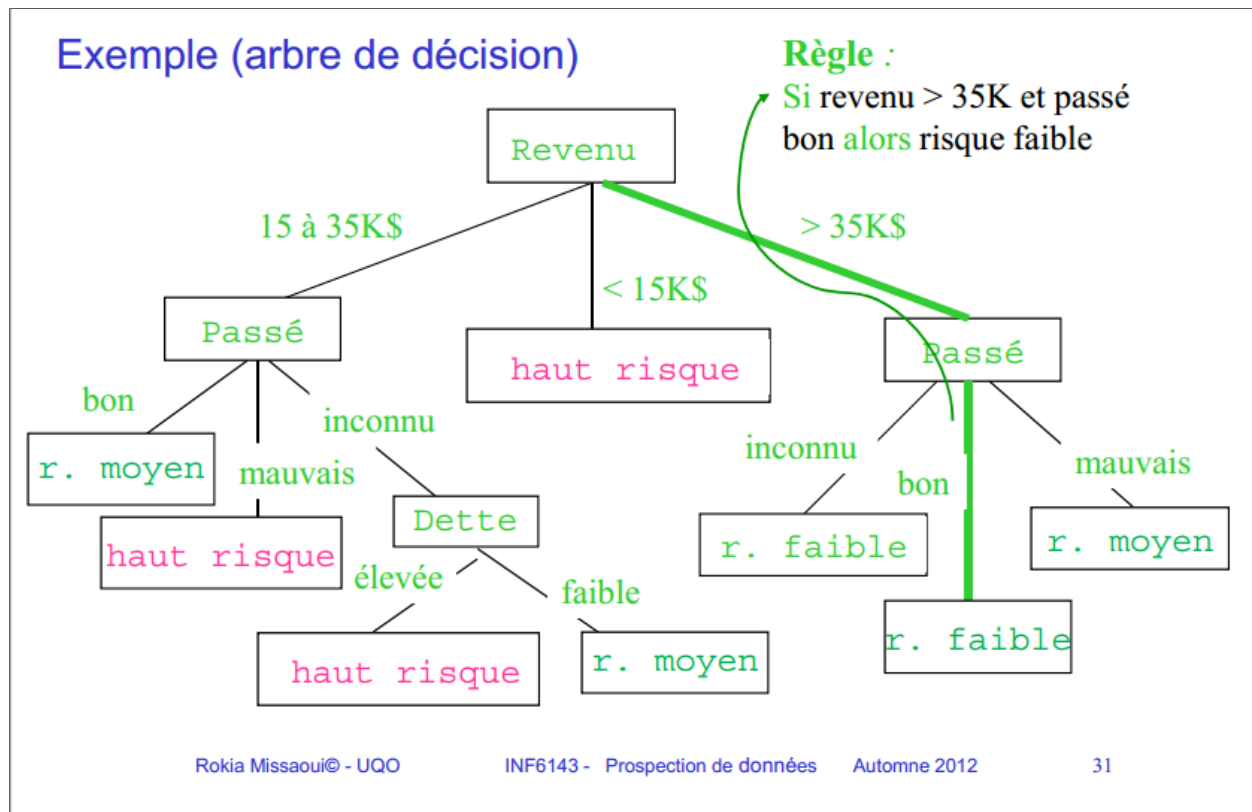


FIG. 5- Arbre de décision pour les demandes de prêts

D'après l'arbre de décision ci-dessus, on peut extraire la règle de classification suivante:

Si revenu < 15K **ou** (15K < = revenu < = 35K **et** passé mauvais) **ou** (15K < = revenu < = 35K **et** passé inconnu **et** dette élevée), **alors** haut risque (de défaut de prêt).

Chapitre 3 Méthodologie

Une fois que les EEG sont enregistrés pour les deux types de participants, l'équipe du laboratoire LNRC manipule et transforme les EEG à l'aide d'un logiciel allemand nommé *Brainvision Analyzer*. Les mesures prises pour analyser les données sont les suivantes :

- 1) Référence des données (Cz ou Mastoïde)
- 2) Filtrage des données pour chaque bande de fréquence
- 3) Correction du signal de mouvements oculaires dans les données
- 4) Retrait de toutes les erreurs engendrées par les mouvements du corps ou de la tête
- 5) Utilisation de l'algorithme de transformée de Fourier rapide qui calcule la puissance des potentiels électriques des électrodes pour les différentes bandes de fréquences.

Ces étapes constituent la phase de prétraitement de données. Tel qu'indiqué auparavant, le processus de découverte de connaissances inclut principalement le prétraitement des données, la fouille de données et l'interprétation des résultats. De notre côté, afin de pouvoir mener notre fouille sur les données EEG, il fallait les faire passer par une seconde étape de prétraitement avant leur utilisation par des modèles de fouille de données.

Dans ce qui suit, nous décrivons en particulier les deux premières étapes sur la base des objectifs établis avec les cliniciens. Dans la section 4, nous présentons et interprétons les résultats préliminaires obtenus pour enchaîner en section 5 avec le deuxième volet d'analyse de données portant principalement sur la fouille au niveau des données transformées des ondes thêta et alpha.

3.1 Prétraitement des données

Les données qui nous ont été fournies par le laboratoire LNRC comprennent différentes bandes (alpha, thêta, delta et bêta) qui se trouvent au niveau de l'EEG et à des moments différents du traitement. La première session (*baseline*) correspond à la phase de prétraitement du patient alors que la deuxième session correspond à des données recueillies une semaine après le début du traitement. Ces bandes représentent des oscillations neuronales dans le système nerveux central enregistrées à partir de l'activité électrique au niveau du cuir chevelu. L'EEG est généralement enregistré à partir de 16 électrodes (sites) ou plus installés au niveau du cuir chevelu du patient. Les données sont ensuite enregistrées avec les yeux fermés puis ouverts des patients. L'activité EEG sur les électrodes est enregistrée par rapport à un site de référence qui peut être soit l'électrode mastoïde, électrode placée en-dessous de l'oreille gauche sur l'os, ou Cz, électrode placée au centre du crâne.

Dans le but de pouvoir mieux exploiter les données qui nous ont été fournies, il a été nécessaire de faire subir à ces données une étape supplémentaire de prétraitement qui consiste à créer de nouveaux ensembles de données qui peuvent être exploités par les algorithmes de fouille de données. Ainsi, les données ont été divisées selon les sessions 1 ou 2, la condition « yeux ouverts » (EO) ou fermés (EC), l'électrode de référence (mastoïde ou Cz) et finalement en deux groupes d'individus : les 52 patients dépressifs et les 43 sujets sains.

En outre, puisque les valeurs de l'EEG (préalablement normalisées) sont continues, il fallait les discrétiser (chaque fois que l'algorithme de fouille de données l'exige) en choisissant l'une des méthodes de discrétisation fournies par le logiciel *Analysis Services de MS SQL Server*.

Après la phase de prétraitement, nous avons constitué trois groupes de données. Le premier groupe contient des données mixtes (ensemble de données avec les caractéristiques et les valeurs EEG pour les patients répondeurs et non-répondeurs). Le deuxième groupe contient deux sous-ensembles de données avec les caractéristiques et les valeurs EEG des patients répondeurs et non répondeurs pris séparément. Le troisième groupe concerne les caractéristiques et les valeurs EEG pour les sujets sains.

Ces différents groupes issus de la phase de prétraitement sont utilisés par les algorithmes retenus de fouille de données.

Il est important de noter qu'en raison du grand nombre d'attributs que peut contenir un ensemble de données, *Analysis Services* nous permet d'effectuer une réduction de la dimensionnalité par la sélection d'attributs (*feature selection*) les plus pertinents pour l'analyse. Cette sélection est essentielle pour une analyse efficace des données et la production de modèles de fouille de données axés sur les attributs les plus importants. Par ailleurs, même si la population d'individus retenus dans l'analyse est relativement faible, cela semble être le cas de la plupart des études reliées à notre problématique à cause principalement de la difficulté à pouvoir mener des collectes de données à plus large échelle et à attirer un plus grand nombre de participants.

3.2 Démarche de fouille de données

À ce jour, les cliniciens du laboratoire LNRC ne peuvent prédire la réponse d'un patient à un traitement antidépresseur donné. Le patient est donc appelé à suivre le traitement prescrit pour une durée de douze semaines même si les soins s'avèrent plus tard inadéquats. Ils collectent les EEG et des mesures cliniques de dépression d'un patient avant le démarrage du traitement et après la fin du traitement. Ces mesures cliniques de dépression (HAM-D, MADRS) sont des questionnaires diagnostiques utilisés par les cliniciens pour mesurer la sévérité de la dépression des patients. Un patient est dit répondeur à un traitement s'il démontre une réduction de 50% dans son score MADRS entre le début et la fin du traitement. Ainsi, une méthode qui permet de

prédire la réponse à un traitement d'une semaine (au lieu de douze semaines) après le début de ce dernier serait d'une grande utilité pour les cliniciens. Ces derniers désirent que l'on se penche sur deux aspects : la prédiction de l'efficacité du traitement et le diagnostic pour discerner entre les patients dépressifs et les sujets sains. L'aspect diagnostique vise alors à définir les caractéristiques spécifiques d'un profil dépressif d'un cerveau donné en comparant l'EEG des patients dépressifs avec celui des sujets sains.

L'approche analytique suggérée dans un premier temps par nos interlocuteurs est d'essayer de voir quelles combinaisons de bandes de fréquences de 28 électrodes, yeux fermés (EC) ou ouverts (EO), et points de référence (mastoïde ou Cz) seraient les meilleures pour discriminer entre les patients répondeurs et non répondeurs (aspect prédiction) d'une part, et entre les patients dépressifs et les sujets sains (aspect diagnostique) d'autre part. D'après les cliniciens, cette approche semblerait prometteuse car c'est la première fois qu'une telle expérimentation serait effectuée avec les données de leurs patients en tenant compte de plusieurs combinaisons possibles de paramètres.

Il existe deux approches pour effectuer une fouille ou analyse de données. La première consiste à émettre des hypothèses en premier lieu puis essayer de les confirmer ou infirmer en analysant les résultats obtenus. La deuxième approche consiste à mener une fouille à l'aveuglette (*trial and error*) ou une fouille exhaustive pour la recherche de connaissances intéressantes en modifiant les valeurs des paramètres des algorithmes utilisés et en alimentant les algorithmes avec des combinaisons de données différentes.

Plusieurs suggestions ont été émises par nos interlocuteurs au laboratoire LNRC durant cette étude. Nous retenons deux d'entre elles. La première proposition consiste à faire une comparaison entre les valeurs des EEG des patients avant et après la deuxième semaine du traitement en vue de détecter si le changement d'activité du EEG peut être utilisé comme le meilleur prédicteur de la réponse au traitement en fonction des bandes (alpha, thêta, delta et bêta) ou d'une combinaison bandes/électrodes (sites). La deuxième suggestion consiste à étudier les valeurs de l'EEG au niveau des électrodes qui sont situées seulement au niveau de deux parties (cf. figure 2) : frontale (F3, F4) et postérieure (P3, P4) du cerveau. Selon la littérature, les deux hémisphères (gauche et droit) servent à différentes fonctions émotionnelles et il serait donc intéressant d'étudier les différences des valeurs EEG au niveau de ces deux parties. Cela peut se faire en calculant la différence entre F4 et F3 et celle entre P4 et P3.

Ces propositions nous ont servi comme point de départ pour nos expérimentations et nous ont aussi inspiré pour trouver d'autres approches d'analyse et d'autres combinaisons de paramètres qui pourraient éventuellement conduire à des résultats intéressants.

Puisque notre recherche se penche sur deux aspects (prédiction et diagnostic), nous allons utiliser d'abord et avant tout la technique de classification. Pour le premier aspect de prédiction, il s'agit de prédire la classe d'appartenance d'un patient (répondeur ou non-répondeur au traitement), sur la base des valeurs de l'EEG recueillies avant et une semaine après le début du traitement. Pour le deuxième aspect, il s'agit de proposer un diagnostic pour une personne sur la base des données de l'EEG collectées à l'arrivée d'un patient. Le modèle de regroupement quant à lui, nous permettra d'identifier des groupes de participants et d'en dégager les caractéristiques. De plus, il est possible que des regroupements inattendus de patients présentant des caractéristiques similaires se dégagent, ce qui peut présenter des informations nouvelles et utiles aux cliniciens. Les règles d'association, quant à elles, nous permettront de découvrir des associations ou liens possibles entre les valeurs EEG des patients et leur réponse finale aux traitements.

Dans les deux prochains chapitres, nous décrivant les tâches de fouille de données que nous avons effectuées en deux étapes selon les recommandations de nos collaborateurs au LNRC et sur la base des travaux de recherche récents en matière d'exploitation des EEG pour les besoins de prédiction. La première étape a concerné l'analyse de diverses ondes cérébrales et en particulier les ondes alpha aussi bien pour le diagnostic que pour la prédiction précoce de l'efficacité d'un traitement alors que la deuxième partie a été dédiée aux ondes thêta pour la prédiction précoce seulement.

Chapitre 4 Fouille de données – Partie 1

On présente ci-dessous un échantillon des résultats obtenus à l'aide des modèles de regroupement, règles d'association, réseaux de neurones et arbres de décision en utilisant le module de services d'analyse (*Analysis Services*) du système *MS SQL Server 2008* et l'outil d'exploration de données XLSTAT de *Microsoft Excel*. Le premier permet d'effectuer diverses tâches en fouille et entreposage de données alors que le second offre une variété de fonctions et d'outils de fouille de données pour améliorer les capacités d'analyse du logiciel Excel de *Microsoft Office*.

4.1 Prédiction

L'ensemble de données nommé *Alpha Tot S2 Ec Cz* qui a été utilisé pour l'obtention de ces résultats contient des données mixtes (caractéristiques et valeurs EEG pour les patients répondeurs et non répondeurs). Il s'agit des valeurs EEG de la bande de fréquence alpha total (8-14 Hz) en référence à l'électrode Cz, pendant que les patients avaient leurs yeux fermés (EC) et une semaine après le début du traitement (S2). Pour un premier essai, nous avons pris en considération une des suggestions proposées par l'équipe du laboratoire LNRC en choisissant seulement quatre électrodes (F3, F4, P3, et P4) pour la construction des modèles. Comme indiqué précédemment, nous avons utilisé le composant *Analysis Services* du logiciel *Microsoft SQL Server* pour conduire la fouille de données selon les trois principales tâches retenues. Il s'agit de la classification par réseaux de neurones, le regroupement par la méthode de *K-means* et la production de règles d'association.

Même si nous avons mené plusieurs tests et obtenu plusieurs résultats, nous fournissons dans ce rapport uniquement les résultats relatifs à l'objectif de prédiction de l'efficacité d'un traitement.

4.1.1 Réseaux de neurones

Analysis Services nous permet de visualiser les différents modèles afin d'obtenir une interprétation claire des résultats.

Le modèle de réseaux de neurones est affiché à l'aide d'une visionneuse qui comporte trois volets : entrée, sortie et variables. Afin de pouvoir construire le modèle, il faut avoir au moins un attribut d'entrée (variable explicative), et un attribut prévisible (variable à expliquer/prédire). Dans notre cas, les attributs d'entrée sont les valeurs EEG recueillies au niveau des quatre

électrodes (F3, F4, P3, et P4). L'attribut prévisible est l'attribut nommé 'répondeur' qui prend la valeur 1 ou 0 indiquant si le patient répond ou ne répond pas au traitement prescrit. Une fois les variables d'entrée et l'attribut prévisible choisis, le modèle associe chaque état possible des attributs d'entrée à chaque état possible de l'attribut prévisible et calcule la probabilité d'occurrence de l'attribut prévisible pour chaque valeur de l'attribut d'entrée.

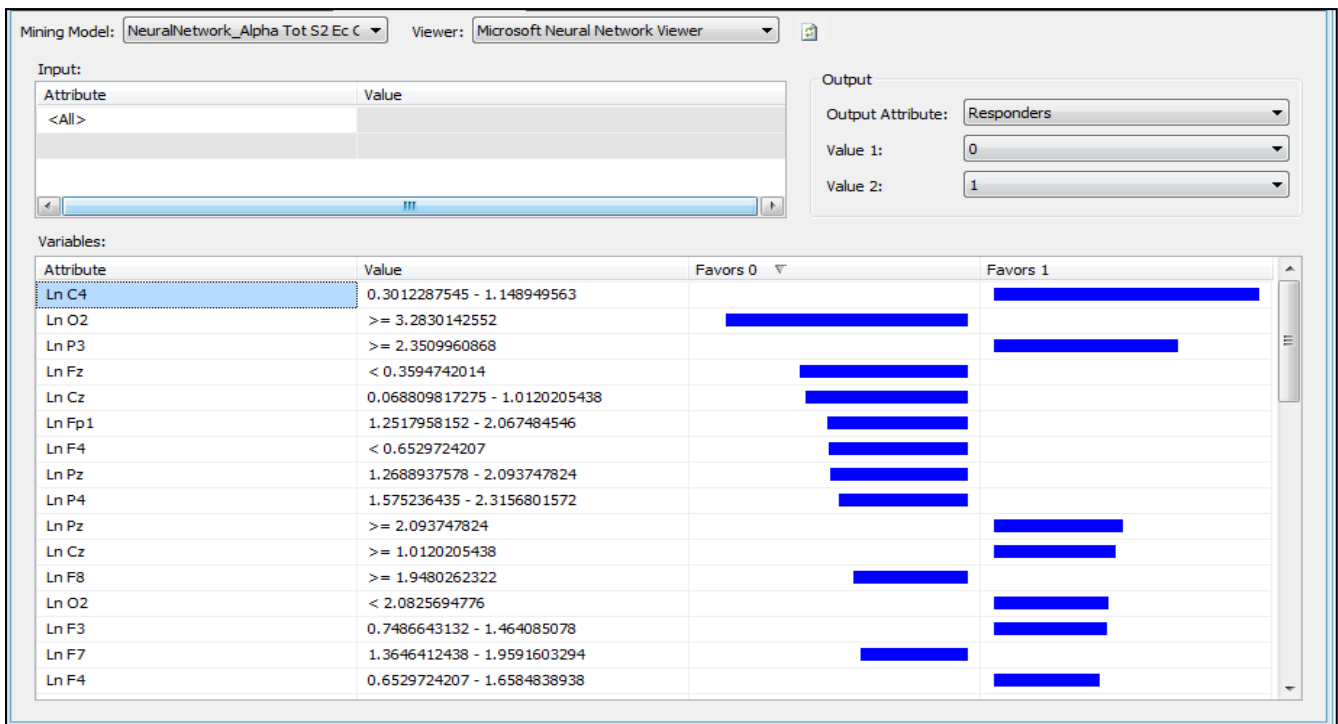


FIG. 6 – Capture d'écran du modèle de réseaux de neurones

Par exemple, on peut voir dans la figure 6 que l'électrode C4 (et plus précisément le log népérien de la valeur de C4) a un score de 100% (indiqué par la taille de la barre) alors que l'électrode F4 a moins d'impact. Comme expliqué auparavant, le score indique l'importance de l'utilisation d'un attribut pour la prédiction du résultat de l'attribut prévisible. Le tableau 2 ci-dessous présente avec plus de détails un échantillon des résultats obtenus après un clic sur les différentes barres bleues. D'après ce tableau, on peut voir que lorsque les valeurs EEG (en logarithme népérien \ln) recueillies au niveau de l'électrode C4 sont comprises entre approximativement 0.3012 et 1.1489, il y a une probabilité de 24.14% que le patient soit non répondeur au traitement (*favors 0*) et une probabilité de 72.29% qu'il soit répondeur au traitement (*favors 1*).

Attribute	Value	Favors 0	Favors 1
Ln C4	0.3012287545 - 1.148949563		Score: 100 Probability of Value1: 24.14 % Probability of Value2: 72.29 %
Ln O2	>= 3.2830142552	Score: 91.32 Probability of Value1: 64.38 % Probability of Value2: 32.05 %	
Ln P3	>= 2.3509960868		Score: 69.85 Probability of Value1: 29.61 % Probability of Value2: 66.82 %
Ln Fz	< 0.3594742014	Score: 63.56 Probability of Value1: 58.59 % Probability of Value2: 37.84 %	
Ln Cz	0.068809817275 - 1.0120205438	Score: 61.58 Probability of Value1: 58.16 % Probability of Value2: 38.27 %	

Tableau 2 - Résultats générés par l'algorithme de réseaux de neurones

4.1.2 Règles d'association

La capture d'écran ci-dessous illustre les règles d'association extraites à partir de l'ensemble de données *Alpha Tot S2 Ec Cz*. La visionneuse est divisée en trois colonnes : probabilité, importance et règle. La probabilité d'une règle, généralement appelée confiance, est une mesure qui détermine la probabilité de la conclusion de la règle quand la prémisse est vraie. Elle prend une valeur comprise entre 0 et 1. L'importance quant à elle, est une mesure de l'utilité d'une règle et représente l'intérêt (*lift*) défini auparavant. Une plus grande valeur signifie une meilleure règle. Finalement, la colonne « Règle » contient les règles extraites de l'ensemble de données *Alpha Tot S2 Ec Cz*.

Puisque dans cette partie nous nous intéressons à la prédiction, nous allons mettre l'accent sur les règles qui comportent dans la partie droite (conclusion) l'attribut de classification répondeur (*Responder*).

Analysis services nous donne la possibilité de modifier les paramètres de l'algorithme de règles d'association pour filtrer les règles de sorte que seules les règles les plus intéressantes sont visibles. Dans cet échantillon, nous avons réglé le seuil minimal de la probabilité à 40% et celui de l'importance à 12%.

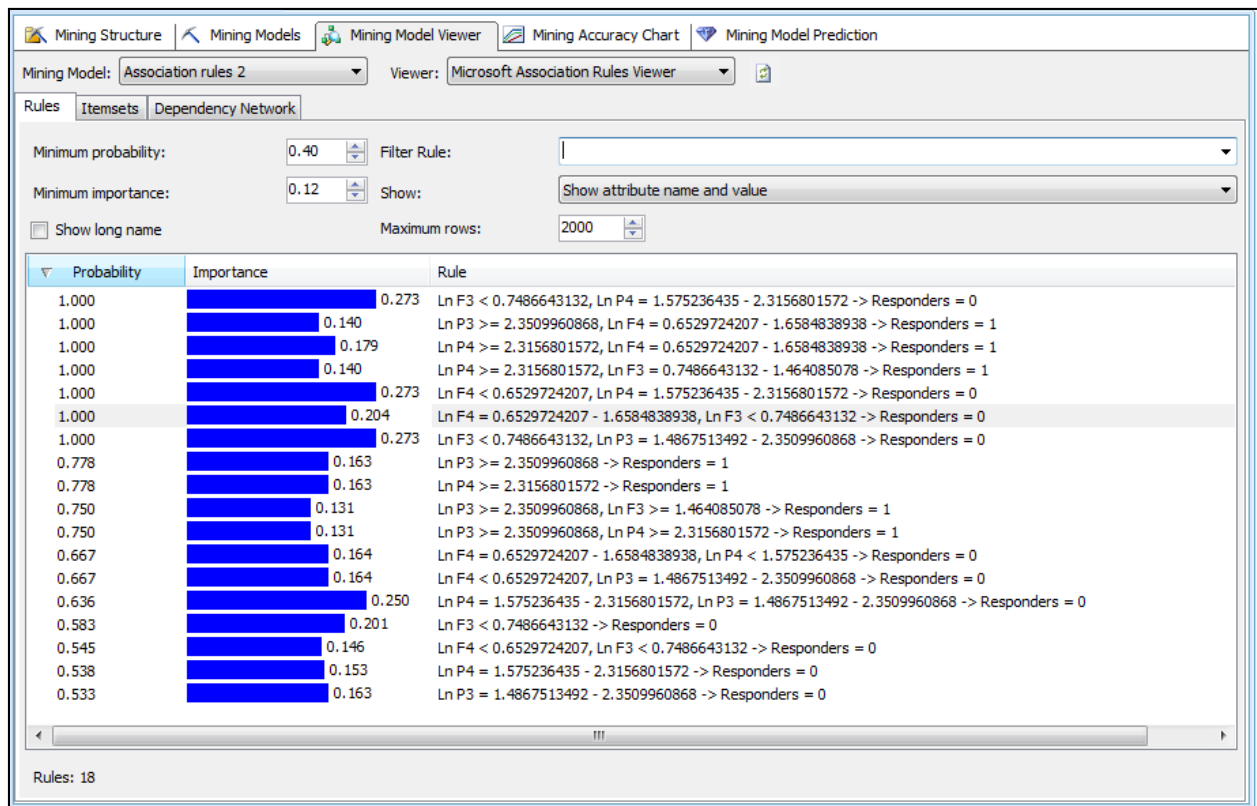


FIG. 7 – Capture d'écran du modèle de règles d'association

Le tableau suivant fournit les règles les plus pertinentes. En prenant la première règle comme illustration, on peut voir que dans le cas où la valeur (en log népérien) EEG recueillie à l'électrode F3 est inférieure à près de 0.749 et la valeur EEG recueillie à l'électrode P4 est approximativement entre 1.575 et 2.315, la probabilité que le patient ne réponde pas favorablement au traitement est de 1.0 (100%).

Probabilité	Importance	Règle
1	0.273001272	Ln F3 < 0.7486643132, Ln P4 = 1.575236435 - 2.3156801572 -> Répondeur = 0
1	0.140375707	Ln P3 >= 2.3509960868, Ln F4 = 0.6529724207 - 1.6584838938 -> Répondeur = 1
1	0.179296399	Ln P4 >= 2.3156801572, Ln F4 = 0.6529724207 - 1.6584838938 -> Responders = 1
1	0.140375707	Ln P4 >= 2.3156801572, Ln F3 = 0.7486643132 - 1.464085078 -> Responders = 1
1	0.273001272	Ln F4 < 0.6529724207, Ln P4 = 1.575236435 - 2.3156801572 -> Responders = 0
1	0.204119983	Ln F4 = 0.6529724207 - 1.6584838938, Ln F3 < 0.7486643132 -> Responders = 0
1	0.273001272	Ln F3 < 0.7486643132, Ln P3 = 1.4867513492 - 2.3509960868 -> Responders = 0
0.778	0.162727297	Ln P3 >= 2.3509960868 -> Responders = 1
0.778	0.162727297	Ln P4 >= 2.3156801572 -> Responders = 1
0.75	0.131404779	Ln P3 >= 2.3509960868, Ln F3 >= 1.464085078 -> Responders = 1
0.75	0.131404779	Ln P3 >= 2.3509960868, Ln P4 >= 2.3156801572 -> Responders = 1
0.667	0.163502132	Ln F4 = 0.6529724207 - 1.6584838938, Ln P4 < 1.575236435 -> Responders = 0
0.667	0.163502132	Ln F4 < 0.6529724207, Ln P3 = 1.4867513492 - 2.3509960868 -> Responders = 0
0.583	0.200659451	Ln F3 < 0.7486643132 -> Responders = 0
0.545	0.146128036	Ln F4 < 0.6529724207, Ln F3 < 0.7486643132 -> Responders = 0
0.538	0.15296746	Ln P4 = 1.575236435 - 2.3156801572 -> Responders = 0
0.533	0.163126282	Ln P3 = 1.4867513492 - 2.3509960868 -> Responders = 0

Tableau 3 – Règles d'association extraites à partir de l'ensemble de données *Alpha Tot S2 Ec Cz*

4.1.3 Regroupement

Comme mentionné précédemment, la méthode de regroupement utilisée dans notre recherche est la méthode de *K-means*. Sans modifier les paramètres de l'algorithme, quatre regroupements ont été déduits comme le montre la figure 8.

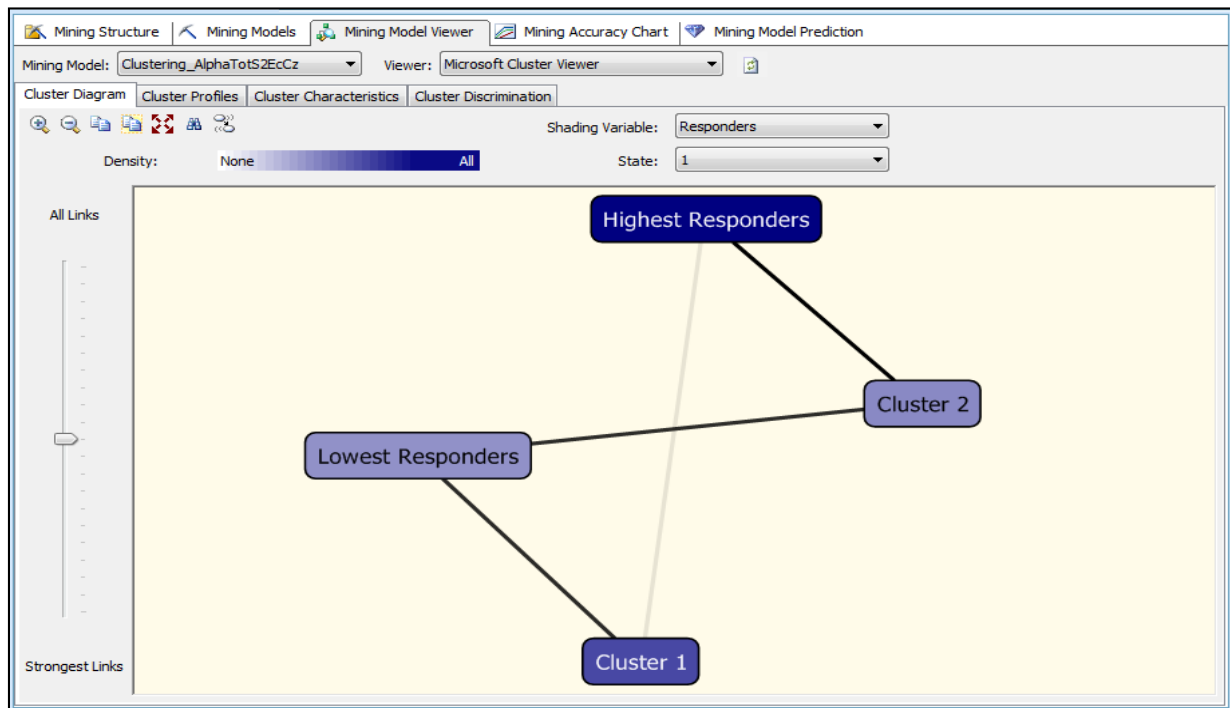


FIG. 8– Capture d'écran du modèle de regroupement

Une consultation des caractéristiques de chacun des groupes nous a permis d'associer une étiquette à deux groupes importants. Le groupe contenant le plus grand pourcentage de patients répondeurs, indiqué par une couleur mauve très foncée, a été étiqueté par *Highest responders*. Celui contenant le plus petit pourcentage de répondeurs est marqué *Lowest responders*. Cependant, puisque le nombre de groupes est grand, la population de chaque groupe est faible. Pour avoir de meilleurs résultats, il est possible d'ajuster les paramètres de l'algorithme pour n'avoir que deux ou trois groupes.

En analysant les groupes générés, nous avons la possibilité de comparer la répartition des valeurs EEG au niveau des électrodes des patients répondeurs et ceux non répondeurs.

D'après la figure 9, on peut voir que les patients appartenant aux groupes *Highest responders* et *Lowest responders* ont tous la valeur EEG au niveau de l'électrode F3 entre environ 0.748 – 1.464. Par contre, les valeurs EEG au niveau de l'électrode F4 sont différentes pour les patients des deux groupes. La plupart des patients appartenant au premier groupe (cf. cluster 2) ont une valeur de Ln F4 < 0.653 tandis que ceux appartenant au deuxième groupe (cf. cluster 3) ont systématiquement une valeur comprise entre approximativement 0.653 et 1.658. Ainsi donc, l'électrode F4 de la bande de fréquence alpha total (8-14 Hz) joue un rôle important qui permet de discriminer les deux groupes considérés.

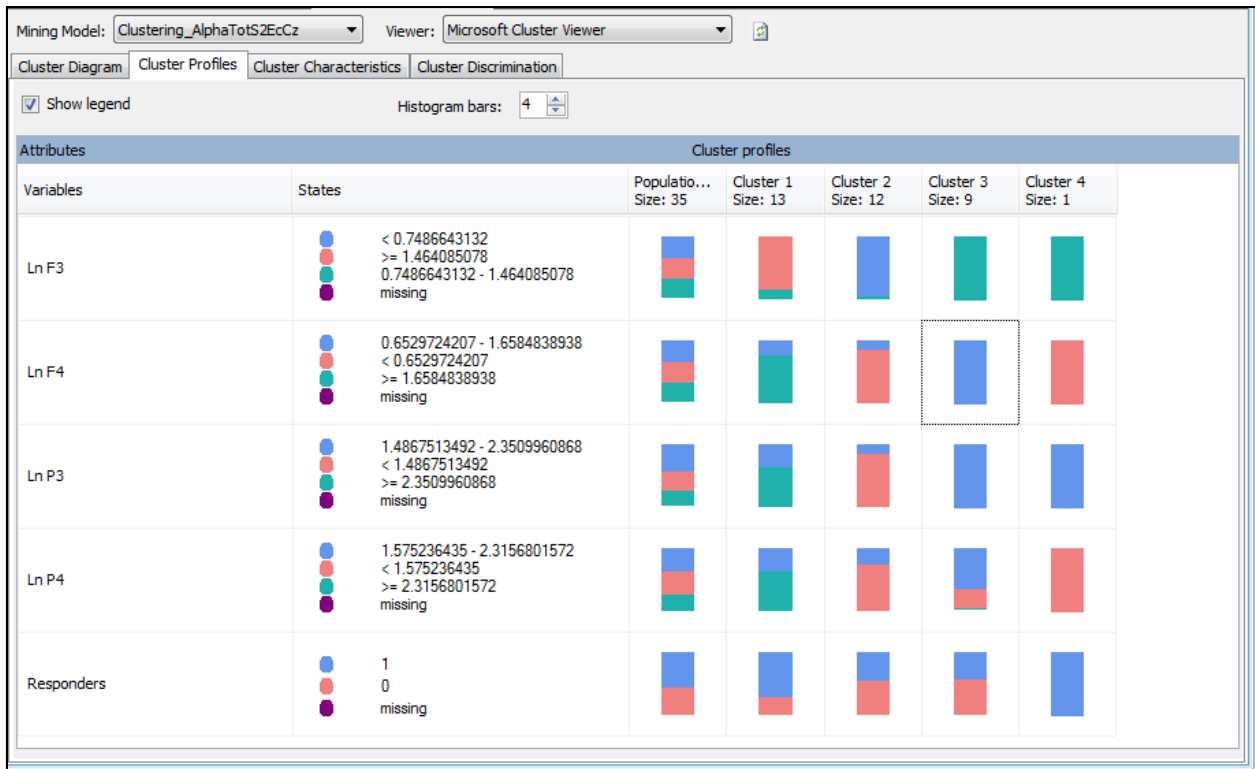


FIG. 9 – Caractéristiques des groupes

4.2 Diagnostic de la dépression

4.2.1 Règles de classification

En ce qui concerne l'aspect diagnostique, les cliniciens du laboratoire nous ont suggéré d'analyser toutes les valeurs EEG des bandes de fréquences alpha ensemble (alpha total, alpha 1 et alpha 2). La raison pour laquelle l'étude s'est limitée à l'analyse de ces bandes est que les données des autres bandes ne nous étaient pas encore disponibles au moment de l'étude. Les données

utilisées étaient celles de la phase prétraitement (session 1), en référence à l'électrode Cz, et pendant que les patients avaient leurs yeux fermés et ouverts (EC et EO).

Nous rappelons que le but de l'aspect diagnostique est de comparer les données EEG des deux groupes d'individus (malades dépressifs et sujets sains) pour identifier des règles de classification qui expliquent le mieux l'état dépressif et qui permettent de prédire la classe d'appartenance d'un individu (sain ou dépressif). Nous avons fait appel aux règles d'association pour identifier les règles de classification car le module *Analysis Services* ne nous a pas permis d'obtenir un arbre de décision à partir des données considérées. Les règles retenues sont celles qui ont l'attribut *Depressed* au niveau de la conclusion et qui ont des mesures de qualité (probabilité et importance) relativement élevées.

La capture d'écran ci-dessous illustre les règles d'association obtenues avec une confiance d'au moins 45%.

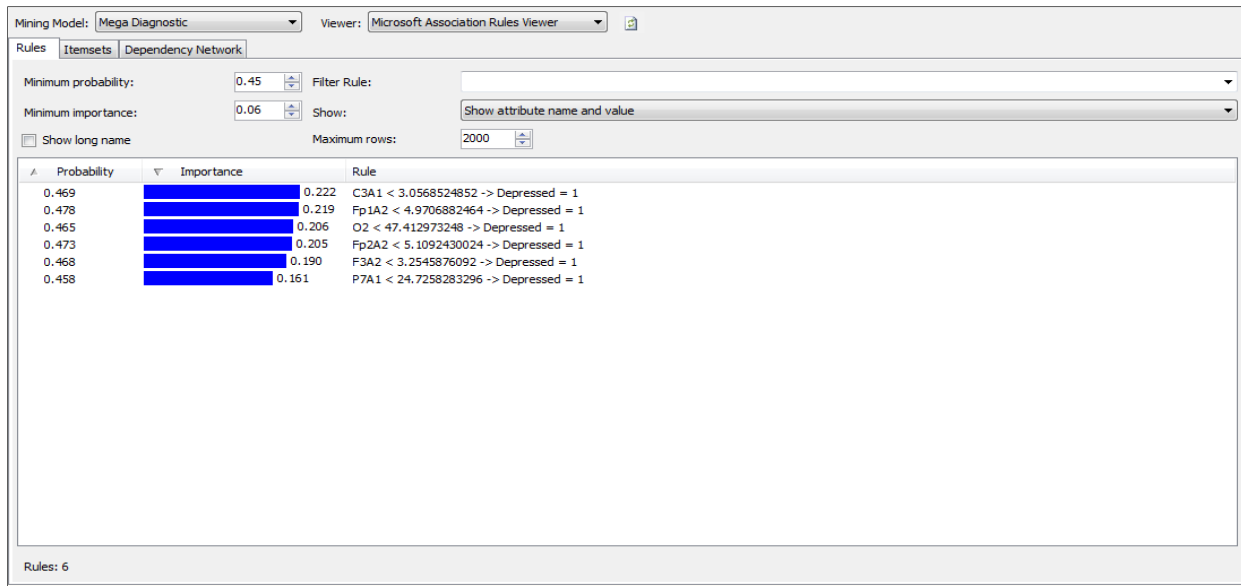


FIG. 10 – Capture d'écran du modèle de règles d'association

Le tableau suivant présente d'une manière plus lisible les règles affichées dans la capture d'écran précédente.

Probability	Importance	Rule
0.469	0.22176978	C3A1 < 3.0568524852 -> Depressed = 1
0.478	0.219353822	Fp1A2 < 4.9706882464 -> Depressed = 1
0.465	0.206167396	O2 < 47.412973248 -> Depressed = 1
0.473	0.204998233	Fp2A2 < 5.1092430024 -> Depressed = 1
0.468	0.190466384	F3A2 < 3.2545876092 -> Depressed = 1
0.458	0.160826942	P7A1 < 24.7258283296 -> Depressed = 1

On remarque que la probabilité et l'importance de ces règles sont trop faibles pour être valablement exploitées pour la prédiction de la classe d'appartenance d'un individu (sain ou dépressif). En prenant la première règle comme exemple, on peut voir que lorsque la valeur EEG recueillie au niveau de l'électrode C3 de la bande de fréquence alpha 1 (A1) est inférieure ou égale à approximativement 3.05, la personne est dépressive. Cette règle est vraie avec une probabilité de 47% et une importance de 22%.

4.2.2 Réseaux de neurones

Nous avons également analysé cet ensemble de données avec le modèle de réseaux de neurones. Le tableau suivant présente les règles obtenues.

Attribute	Value	Favors 0	Favors 1
CP5A2	>= 26.2997505184		Score: 100
			Probability of Value1: 21.39 %
			Probability of Value2: 77.59 %
FC2A2	>= 8.72295828		Score: 97.09
			Probability of Value1: 21.96 %
			Probability of Value2: 77.02 %
C4A1	>= 28.2085095744		Score: 92.78
			Probability of Value1: 22.84 %
			Probability of Value2: 76.14 %
F3	>= 57.7063512896	Score: 83.1	
		Probability of Value1: 69.38 %	
		Probability of Value2: 29.60 %	
FC1	>= 29.3249999968	Score: 81.43	
		Probability of Value1: 68.97 %	
		Probability of Value2: 30.01 %	
FC6A1	>= 41.6666666624		Score: 74.49
			Probability of Value1: 26.81 %
			Probability of Value2: 72.17 %
CP2A2	6.9266211 - 9.1192434	Score: 73.75	
		Probability of Value1: 67.06 %	
		Probability of Value2: 31.92 %	

La première règle du tableau précédent a un score de 100% et indique que lorsque les valeurs EEG recueillies au niveau de l'électrode CP5 de la bande de fréquence alpha 2 (A2) sont supérieures ou égales à approximativement 26.3, il y a une probabilité de 21.39% que le patient ne soit pas dépressif (*favors 0*) et une probabilité de 77.59% qu'il le soit (*favors 1*).

Chapitre 5 Fouille de données – Partie 2

Pour la deuxième phase de notre projet, les cliniciens nous ont vivement recommandé de nous concentrer sur la prédiction de l'efficacité d'un traitement à une étape précoce et de laisser de côté l'aspect diagnostique. En outre, l'équipe du laboratoire LNRC nous a suggéré de nouvelles approches analytiques en utilisant l'EEG absolue et la « cordance » laquelle intègre à la fois les valeurs absolues et relatives des EEG.

La première approche suggérée était de combiner les données EEG absolues de toutes les bandes de fréquences. À titre d'exemple, au lieu d'utiliser la bande de fréquence alpha seulement, il a été recommandé de faire usage d'une combinaison de deux types de bandes de fréquences parmi alpha, bêta et thêta.

Une deuxième approche consiste à analyser les valeurs EEG de toutes les bandes de fréquences avec les variables cliniques (HAM-D et MADRS) et les changements de score des variables cliniques (H17ChngW1B%, H17ChngW12%, MadChngW1B% et MadChngW12B%), pour voir si on peut obtenir de meilleurs résultats. Ces changements de score font référence au changement clinique dans la dépression des patients tel que mesuré par:

- H17ChngW1B% qui mesure le changement de score *du 17-item Hamilton Depression Rating Scale* (H17) entre la phase prétraitement (**Baseline**) et une semaine (W1) après le début du traitement
- H17ChngW12% qui mesure le changement de H17 entre la phase prétraitement et douze semaines (W12) après le début du traitement
- MadChngW1B% qui mesure le changement de score MADRS entre la phase de prétraitement et une semaine après le début du traitement
- MadChngW12B% qui mesure le changement de score MADRS entre la phase prétraitement et douze semaines après le début du traitement.

Les données fournies ont été exploitées en utilisant d'abord l'outil d'exploration de données XLSTAT de Microsoft Excel pour produire des arbres de décision et donc des règles de classification. Nous avons ensuite utilisé le prototype *Lattice Miner* développé au laboratoire LARIM de l'UQO pour générer des règles d'association avec ou sans négation des variables préalablement discrétisées puis généralisées.

5.1 Classification par arbre de décision

Nous constatons que les règles de classification obtenues à partir de l'arbre de décision produit en utilisant les données EEG de l'ensemble des bandes de fréquences ne concernent que les valeurs EEG de la bande thêta. Cette constatation vient confirmer l'importance de cette bande à discriminer entre les patients répondeurs et non répondeurs ainsi que les résultats d'études antérieures [9, 10, 14]. Les règles d'association sont présentées dans le tableau ci-après.

Probability	Importance	Rule
1	0.745966567	Ln Fz Theta S2 Mast < 1.0506190858, Ln CP1 Theta S2 Mast < 0.8222916209 -> Responders = 0
1	0.745966567	Ln CP2 Theta S2 Mast < 0.7978548677, Ln CP1 Theta S2 Mast < 0.8222916209 -> Responders = 0
1	0.745966567	Ln CP2 Theta S2 Mast < 0.7978548677 -> Responders = 0
1	0.745966567	Ln F3 Theta S2 Mast < 1.1436397208 -> Responders = 0
1	0.745966567	Ln Fz Theta S2 Mast < 1.0506190858 -> Responders = 0
1	0.745966567	Ln Fz Theta S2 Mast < 1.0506190858, Ln FC1 Theta S2 Mast < 0.786177075 -> Responders = 0
1	0.745966567	Ln Fz Theta S2 Mast < 1.0506190858, Ln CP2 Theta S2 Mast < 0.7978548677 -> Responders = 0
1	0.745966567	Ln FC1 Theta S2 Mast < 0.786177075, Ln CP1 Theta S2 Mast < 0.8222916209 -> Responders = 0
1	0.745966567	Ln F4 Theta S2 Mast < 1.1321521872, Ln F3 Theta S2 Mast < 1.1436397208 -> Responders = 0
1	0.745966567	Ln F4 Theta S2 Mast < 1.1321521872 -> Responders = 0
1	0.745966567	Ln FC1 Theta S2 Mast < 0.786177075, Ln CP2 Theta S2 Mast < 0.7978548677 -> Responders = 0

Tableau 4– Règles d’association obtenues en tenant compte de toutes les bandes de fréquences

En prenant la première règle comme illustration, on peut voir que lorsque la valeur EEG recueillie à l’électrode Fz pendant la deuxième session (après une semaine de traitement) en référence avec l’électrode mastoïde est inférieure à près de 1.05 et la valeur EEG recueillie à l’électrode CP1 pendant la deuxième session en référence à l’électrode mastoïde est moins qu’approximativement 0.82, le patient est non répondeur au traitement avec une probabilité de 1.0 (100%).

Comme mentionné précédemment, nous avons également combiné les données EEG de toutes les bandes de fréquences avec les valeurs des variables cliniques (HAM-D, MADRS) et les changements de score des variables cliniques (H17ChngW1B%, H17ChngW12%, MadChngW1B%, et MadChngW12B%). Nous avons ensuite analysé les données EEG de la bande thêta et la bande alpha totale puisqu’elles semblent être les plus discriminantes d’après les cliniciens du laboratoire.

On constate que les règles de classification obtenues avec la bande thêta ne comprennent ni les valeurs des variables cliniques (HAM-D et MADRS) ni les changements de score de ces derniers (H17ChngW1B%, H17ChngW12%, MadChngW1B%, MadChngW12B%). Cela signifie que ces variables n'ont pas autant d'impact que la bande thêta pour classifier les patients dans l'une des deux catégories : répondeurs et non répondeurs.

Le tableau ci-dessous présente les règles de classification obtenues.

Node	Pred(Responders)	Frequency	Purity	Rules
Node1	1.000	20	51.28%	
Node2	0.000	19	95.00%	If InCP1ThetaS2Mast in [3, 5.5[then Responders = 0 in 95% of cases
Node3	1.000	19	100.00%	If InCP1ThetaS2Mast in [5.5, 11[then Responders = 1 in 100% of cases
Node4	0.000	1	50.00%	If InT7ThetaS2Mast in [3, 4.5[and InCP1ThetaS2Mast in [3, 5.5[then Responders = 0 in 50% of cases
Node5	0.000	18	100.00%	If InT7ThetaS2Mast in [4.5, 7[and InCP1ThetaS2Mast in [3, 5.5[then Responders = 0 in 100% of cases
Node6	0.000	1	100.00%	If InF3ThetaS2Mast in [4, 5[and InT7ThetaS2Mast in [3, 4.5[and InCP1ThetaS2Mast in [3, 5.5[then Responders = 0 in 100% of cases
Node7	1.000	1	100.00%	If InF3ThetaS2Mast in [5, 6[and InT7ThetaS2Mast in [3, 4.5[and InCP1ThetaS2Mast in [3, 5.5[then Responders = 1 in 100% of cases

Tableau 5– Règles d'associations obtenues en analysant la bande de fréquence thêta avec les valeurs et les changements de score des variables cliniques

Les règles en gras sont les règles avec le plus haut support. La première règle est vraie pour tous les 19 patients non répondeurs et donc avec un support de 100% (*purity* = 19/19). La deuxième règle est vraie pour 19 des 20 patients répondeurs (*purity* = 19/20= 95%) alors que la troisième est vraie pour 18 des 19 patients non répondeurs. Le tableau ci-dessous présente les règles de classification obtenues avec la bande alpha totale et les données des variables cliniques.

Node	Pred(Responders)	Frequency	Purity	Rules
Node1	1.000	20	51.28%	
Node2	1.000	19	90.48%	If H17ChngW12B% in [1, 3[then Responders = 1 in 90.5% of cases
Node3	0.000	14	93.33%	If H17ChngW12B% in [3, 8[then Responders = 0 in 93.3% of cases
Node4	0.000	3	100.00%	If H17ChngW12B% in [8, 15[then Responders = 0 in 100% of cases
Node5	1.000	2	100.00%	If lnC4AlphaTotS2Cz in [3, 3.5[and H17ChngW12B% in [1, 3[then Responders = 1 in 100% of cases
Node6	0.000	2	50.00%	If lnC4AlphaTotS2Cz in [3.5, 4.5[and H17ChngW12B% in [1, 3[then Responders = 0 in 50% of cases
Node7	1.000	15	100.00%	If lnC4AlphaTotS2Cz in [4.5, 9[and H17ChngW12B% in [1, 3[then Responders = 1 in 100% of cases
Node8	1.000	2	100.00%	If lnF8AlphaTotS2Cz in [6, 6.5[and lnC4AlphaTotS2Cz in [3.5, 4.5[and H17ChngW12B% in [1, 3[then Responders = 1 in 100% of cases

Tableau 6– Règles d’association obtenues en analysant la bande de fréquence alpha totale avec les valeurs et les changements de score des variables cliniques

En analysant les règles portant sur les ondes alpha, on s’aperçoit que seuls les changements de score des variables cliniques entre la phase de prétraitement et douze semaines après les soins ont été utilisés (H17ChngW12%, MadChngW12B%). Ces règles sont inutiles pour la prédiction précoce de la réponse au traitement. En outre, l’absence des données sur la bande alpha pour la session 2 (une semaine après le début du traitement) dans les règles nous permet de conclure que la bande alpha n’a pas d’impact sur la prédiction précoce de la classe d’appartenance des patients.

L’équipe du laboratoire LNRC nous a suggéré de calculer les valeurs moyennes de l’EEG absolue des électrodes de chaque région cérébrale (cf. les figures 3 et 9) afin de déterminer laquelle de

ces régions est la plus corrélée à la réponse positive ou négative au traitement antidépresseur. Les valeurs EEG pour la session 2 (une semaine après le début du traitement) ont été utilisées. Une fois les moyennes calculées, le modèle d'arbre de décision généré par XLSTAT a été utilisé pour identifier les règles de classification.

Le schéma ci-après² donne une interprétation d'un nœud de l'arbre de décision (régression) produit par XLSTAT.

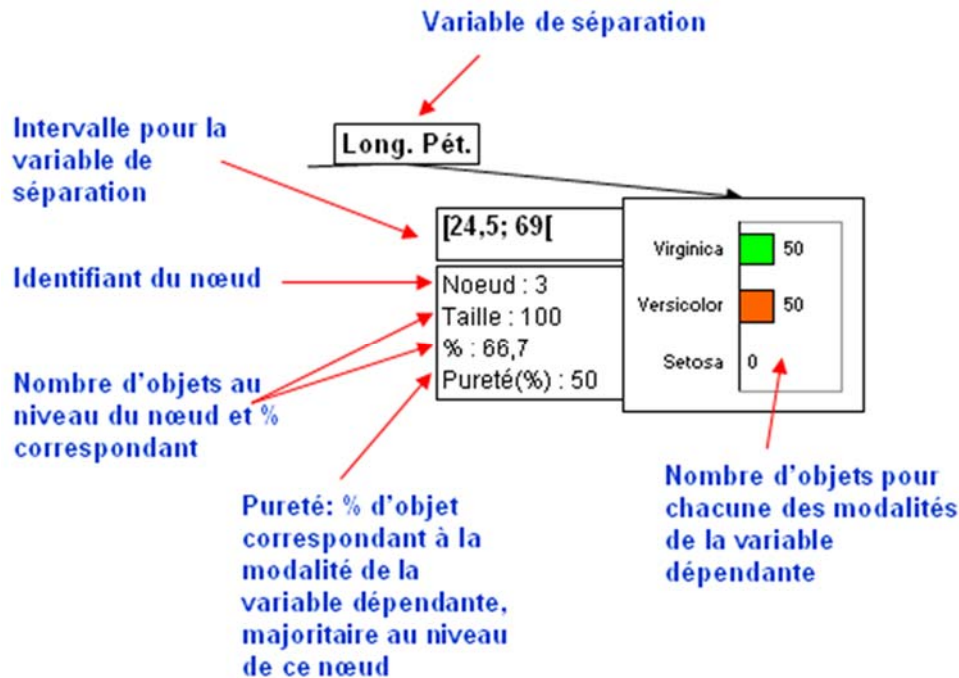


Fig. 10 – Signification du contenu d'un nœud de l'arbre de décision.

La figure suivante (Fig. 11) présente l'arbre de décision généré par XLSTAT à partir des valeurs EEG moyennes des régions cérébrales de la bande thêta. Les abréviations MF et LF désignent respectivement les régions *Midline Frontal* et *Left Frontal* tel qu'indiqué dans le tableau 10.

² Cf. le tutoriel à l'adresse : <http://www.xlstat.com/fr/centre-d-apprentissage/tutoriels/creer-un-arbre-de-classification-chaid-avec-xlstat.html>.

Légende

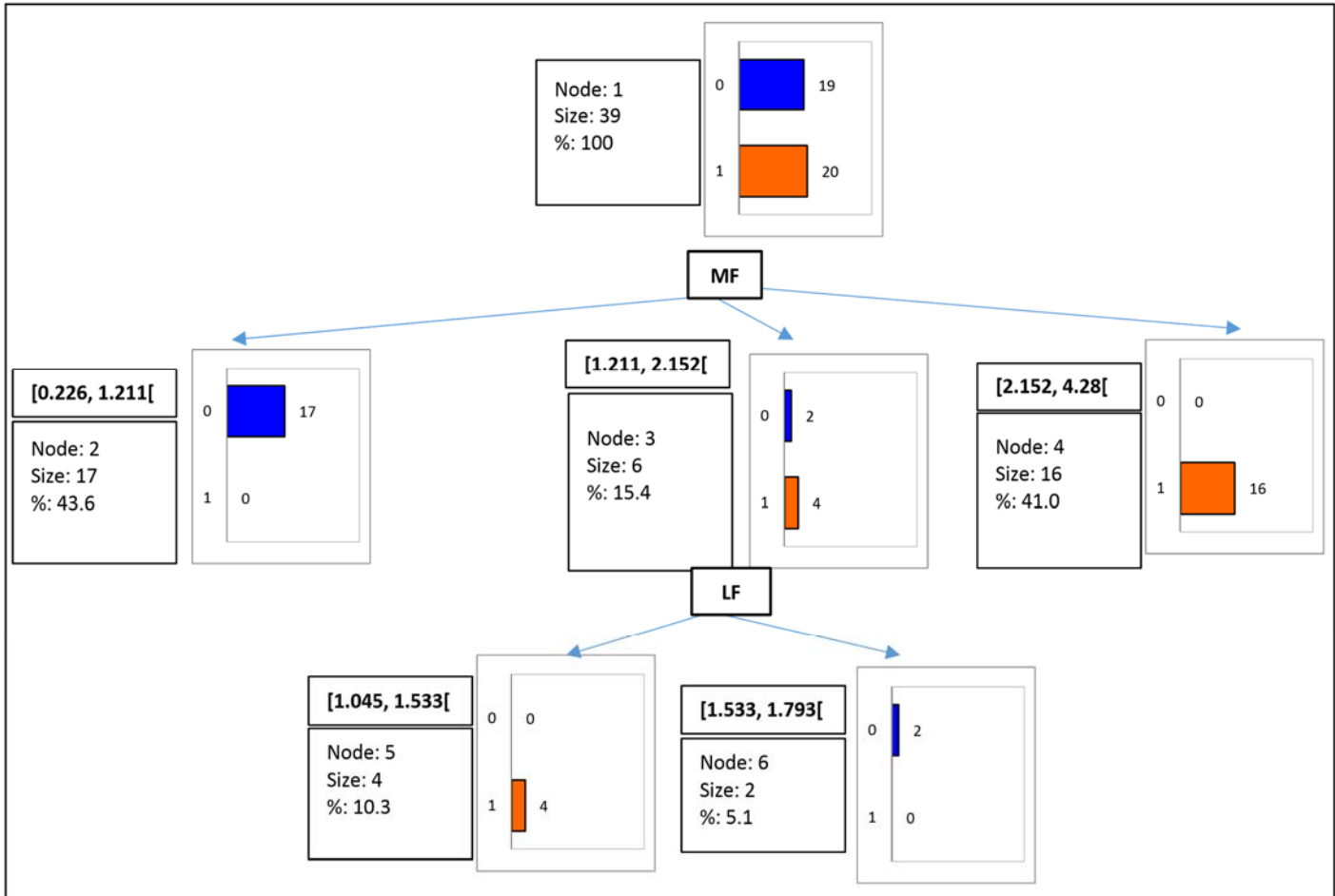
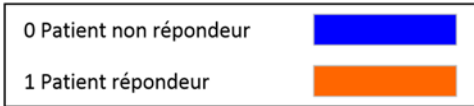


Fig. 11 - Arbre de décision des régions cérébrales de la bande thêta

Le diagramme de la figure 11 permet de visualiser les étapes successives au cours desquelles l'algorithme d'arbre de décision identifie les variables (régions cérébrales) qui permettent de mieux séparer les catégories de la variable dépendante (patients répondeurs ou non répondeurs). On appelle la variable au sommet de l'arbre le *critère* de fractionnement qui en fait correspond à la variable ayant le plus d'impact sur l'attribut de classification (variable dépendante). Ainsi, nous voyons qu'en utilisant la région *Midline Frontal*, l'algorithme a trouvé une règle qui permet de séparer parfaitement les patients répondeurs et non répondeurs. Si la valeur EEG absolue se trouve dans l'intervalle $[0.226, 1.211[$, alors le patient est nécessairement non répondeur. Si par contre elle se trouve dans l'intervalle $[2.152, 4.28[$, alors le patient est répondeur. Le tableau suivant fournit les règles de classification extraites de l'arbre.

Node	Pred(Responders)	Frequency	Rules
Node1	1.000	20	
Node2	0.000	17	If MF in [0.226, 1.211[then Responders = 0 in 100% of cases
Node3	1.000	4	If MF in [1.211, 2.152[then Responders = 1 in 66.7% of cases
Node4	1.000	16	If MF in [2.152, 4.28[then Responders = 1 in 100% of cases
Node5	1.000	4	If LF in [1.045, 1.533[and MF in [1.211, 2.152[then Responders = 1 in 100% of cases
Node6	0.000	2	If LF in [1.533, 1.793[and MF in [1.211, 2.152[then Responders = 0 in 100% of cases

Tableau 7 – Règles de classification des régions cérébrales de la bande thêta

En analysant les règles, on s’aperçoit que toutes les règles concernent les deux régions *Midline Frontal* et *Left Frontal* seulement. Ces résultats confirment ceux énoncés dans des travaux de recherche antérieurs qui indiquent que les régions préfrontales jouent un rôle important dans la discrimination entre les patients répondeurs et non répondeurs [9, 10, 14,15]. Parmi les règles présentées ci-dessus, les règles les plus importantes (indiquées en gras) sont celles qui concernent l’attribut le plus discriminant et couvrent une bonne proportion de cas, soit 17 (sur 19) patients non répondeurs et 16 (sur 20) patients répondeurs.

L’arbre de décision généré pour la bande alpha totale était trop grand pour être présenté dans ce mémoire. Le tableau ci-dessous présente les règles de classification extraites de l’arbre où la signification des variables relatives aux régions cérébrales est fournie dans le tableau 10.

Node	Frequency	Purity	Rules
Node1	20	0.512820513	
Node2	3	1	If LCT in [0.658, 0.998[then Responders = 1 in 100% of cases
Node3	4	0.8	If LCT in [0.998, 1.595[then Responders = 0 in 80% of cases
Node4	4	0.666666667	If LCT in [1.595, 1.879[then Responders = 0 in 66.7% of cases

Node5	5	0.555555556	If LCT in [1.879, 2.239[then Responders = 0 in 55.6% of cases
Node6	4	0.666666667	If LCT in [2.239, 2.565[then Responders = 0 in 66.7% of cases
Node7	4	0.666666667	If LCT in [2.565, 2.992[then Responders = 1 in 66.7% of cases
Node8	4	1	If LCT in [2.992, 3.36[then Responders = 1 in 100% of cases
Node9	1	0.5	If MP in [0.968, 1.152[and LCT in [0.998, 1.595[then Responders = 0 in 50% of cases
Node10	3	1	If MP in [1.152, 1.706[and LCT in [0.998, 1.595[then Responders = 0 in 100% of cases
Node11	1	1	If RF in [1.54, 1.885[and MP in [0.968, 1.152[and LCT in [0.998, 1.595[then Responders = 1 in 100% of cases
Node12	1	1	If RF in [1.885, 2.23[and MP in [0.968, 1.152[and LCT in [0.998, 1.595[then Responders = 0 in 100% of cases
Node13	2	0.666666667	If LP in [1.423, 1.806[and LCT in [1.595, 1.879[then Responders = 1 in 66.7% of cases
Node14	3	1	If LP in [1.806, 2.319[and LCT in [1.595, 1.879[then Responders = 0 in 100% of cases
Node15	1	1	If MP in [1.294, 1.67[and LP in [1.423, 1.806[and LCT in [1.595, 1.879[then Responders = 1 in 100% of cases
Node16	1	0.5	If MP in [1.67, 2.126[and LP in [1.423, 1.806[and LCT in [1.595, 1.879[then Responders = 0 in 50% of cases
Node17	2	1	If MF in [1.084, 1.231[and LCT in [1.879, 2.239[then Responders = 1 in 100% of cases
Node18	1	0.5	If MF in [1.231, 1.56[and LCT in [1.879, 2.239[then Responders = 0 in 50% of cases
Node19	2	0.666666667	If MF in [1.56, 2.008[and LCT in [1.879, 2.239[then Responders = 0 in 66.7% of cases
Node20	2	1	If MF in [2.008, 2.89[and LCT in [1.879, 2.239[then Responders = 0 in 100% of cases

Node21	2	1	If RF in [2.351, 2.665[and LCT in [2.239, 2.565[then Responders = 0 in 100% of cases
Node22	2	0.5	If RF in [2.665, 3.634[and LCT in [2.239, 2.565[then Responders = 0 in 50% of cases
Node23	1	1	If RP in [1.927, 2.34[and MF in [1.56, 2.008[and LCT in [1.879, 2.239[then Responders = 0 in 100% of cases

Tableau 8– Règles de classification des régions cérébrales de la bande alpha total

En analysant ces règles sur les données alpha, on s’aperçoit que leur support est faible contrairement à celles extraites de l’arbre de décision des régions cérébrales préfrontales de la bande thêta. Cela confirme les résultats de travaux antérieurs qui indiquent que la bande de fréquence thêta est la meilleure pour discriminer entre les patients répondeurs et non répondeurs.

Comme mentionné dans l’état de l’art, des études antérieures ont montré que la diminution précoce de la « cordance » de l’EEG préfrontal au niveau de la bande thêta précédait l’amélioration des symptômes de différents médicaments antidépresseurs [8].

Dans une étude menée par Bares et al [9], les valeurs moyennes ont été calculées à partir des trois électrodes frontales (FP1, FP2, Fz) pour chaque patient au niveau de la bande de fréquence thêta. Les résultats de l’étude ont montré que les patients répondeurs démontrent une diminution de la « cordance » de l’EEG préfrontal dès la fin de la première semaine de traitement. La première analyse menée avec les données de la « cordance » visait la réplique des résultats obtenus par Bares et al [9].

Pour cette étape de notre analyse, le laboratoire nous a fourni un jeu de données de 39 patients qui ont subi un traitement de six semaines par deux antidépresseurs. Ces données concernent des valeurs EEG de la mesure de la « cordance » de la bande thêta en vue d’examiner les caractéristiques discriminatoires de cette bande de fréquences. Ces données ont été enregistrées pendant la première session (phase de prétraitement) et la deuxième session du traitement. Il est important de noter que le laboratoire ne nous a fourni que les valeurs EEG de la « cordance » de dix-huit parmi les trente-cinq électrodes et qui sont : F3, Cz, C4, F8, Fz, C3, F4, F7, FP1, FP2, O1, O2, O3, P3, P4, PZ, T3, T4, T5, et T6. Les valeurs moyennes ont été calculées à partir des trois électrodes frontales (FP1, FP2, Fz) pour chaque patient comme cela a été fait dans [9]. Les données ont été identifiées selon la session (1 ou 2) et selon la catégorie de patients (répondeurs ou non répondeurs).

Le modèle de regroupement a été utilisé pour comparer la valeur moyenne de la « cordance » de la première session avec celle de la deuxième session pour les patients répondeurs et non

répondeurs. Les paramètres de l’algorithme ont été modifiés de telle façon que seulement un regroupement soit créé pour chaque session et catégorie de patients.

En effet, les résultats obtenus ont montré une diminution de la valeur moyenne de la « cordance » au niveau de la région préfrontale pour les patients répondeurs et non répondeurs entre la première et la deuxième session (cf. les figures 12 à 14). Cependant, cette diminution était importante chez les patients répondeurs alors qu’il n’y avait pas de changement significatif pour les patients non répondeurs.

La valeur de la « cordance » moyenne pour les patients répondeurs pendant la première session était de 0.17 et a diminué jusqu’à -0.88 lors de la deuxième session. Pour les patients non répondeurs, la valeur de la cordance moyenne pendant la première session était de 0.30 et a diminué jusqu’à 0.15 lors de la deuxième session. Le losange représente la répartition des états de la variable continue « Valeur Moyenne Pref» qui contient la moyenne des valeurs EEG des électrodes préfrontales citées précédemment. Par exemple dans la figure ci-dessous, la valeur minimale de la variable « Valeur Moyenne Pref» est de -1.81, la valeur moyenne est de 0.17 et la valeur maximale est de 2.48.

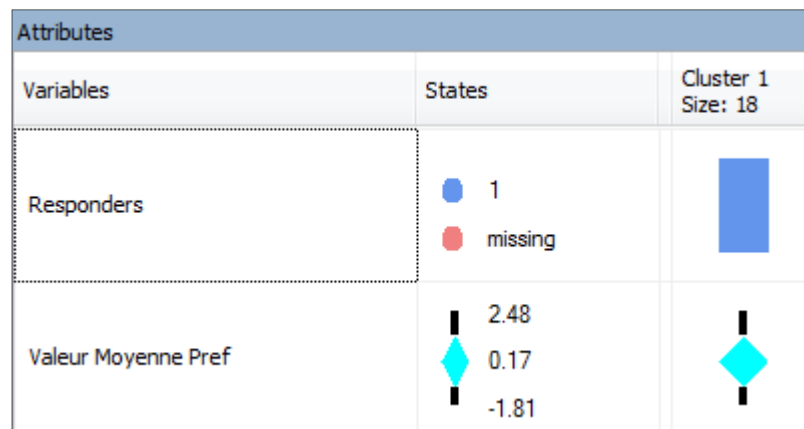


Fig. 12- Répartition de la valeur moyenne de la région préfrontale de la « cordance » chez les patients répondeurs pendant la première session

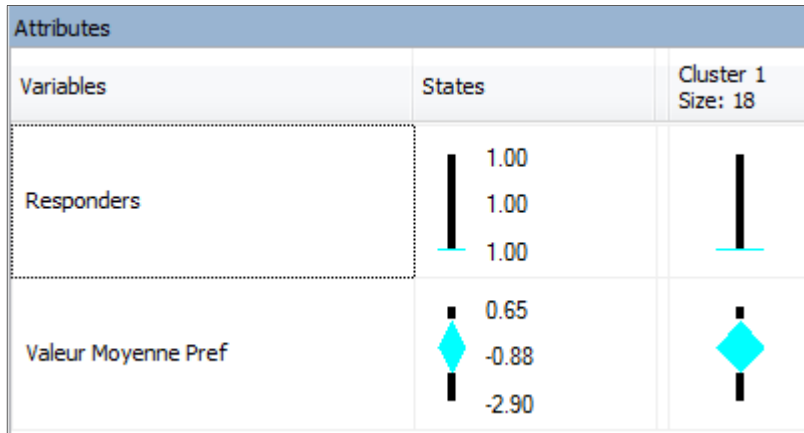


Fig. 13 - Répartition de la valeur moyenne de la région préfrontale de la « cordance » chez les patients répondeurs pendant la deuxième session

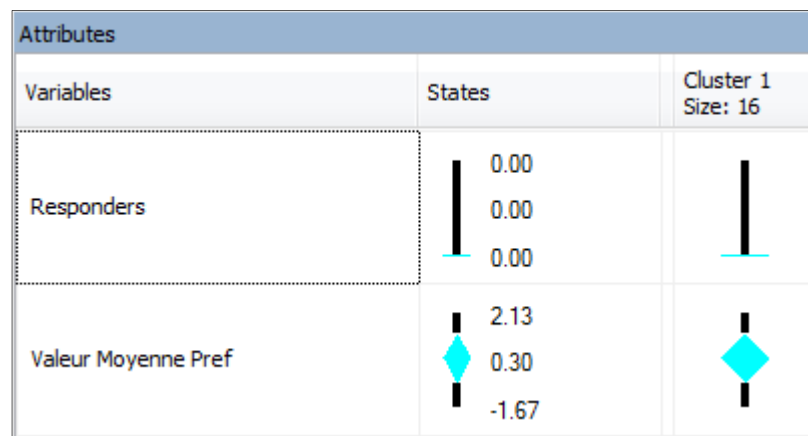


Fig. 14 - Répartition de la valeur moyenne de la région préfrontale de la « cordance » chez les patients non répondeurs pendant la première session




Attributes		
Variables	States	Cluster 1 Size: 16
Responders	<ul style="list-style-type: none"> ● 0 ● missing 	
Valeur Moyenne Pref	<ul style="list-style-type: none"> 2.00  0.15 -1.03 	

Fig. 15 - Répartition de la valeur moyenne de la région préfrontale de la « cordance » chez les patients non répondeurs pendant la deuxième session

La seconde analyse menée avec les données de la « cordance » suite à une suggestion de l'équipe du laboratoire LNRC avait pour but de répliquer les résultats obtenus par Ian A. Cook et al. [10]. La moyenne des valeurs de la cordance de la bande thêta aux électrodes FPZ, FZ, FP2, AF2, F4 et F8 a été calculée pour former la mesure de la « cordance » régionale MRF (*midline and right frontal*) pour chaque patient. Le modèle de regroupement a été utilisé avec les mêmes paramètres modifiés dans l'analyse précédente. Contrairement aux résultats obtenus dans l'analyse précédente, les résultats de cette seconde analyse ont montré une diminution de la valeur moyenne de la « cordance » au niveau de la MRF pour les patients répondeurs seulement. La valeur de la « cordance » moyenne pour les patients répondeurs pendant la première session était de 0.31 et a diminué jusqu'à -0.63 lors de la deuxième session. Pour les patients non répondeurs, la valeur de la « cordance » moyenne pendant la première session était de -0.21 et a augmenté jusqu'à 0.07 lors de la deuxième session.

La troisième analyse suggérée par l'équipe du laboratoire LNRC était de construire un arbre de décision pour comparer nos résultats avec ceux cités par Rabinoff et al. [14]. Les différences des valeurs EEG de la bande thêta entre la première et la deuxième session ont été calculées comme cela a été fait dans [14]. L'arbre de décision généré est trop grand pour être présenté dans ce document mais le tableau ci-dessous illustre quelques-unes des règles de classification extraites de l'arbre.

Node	Pred(Responders)	Frequency	Rules
Node1	1.000	25	
Node2	0.000	4	If f3d in [-2.514, -1.793[then Responders = 0 in 100% of cases
Node3	1.000	4	If f3d in [-1.793, -1.323[then Responders = 1 in 100% of cases
Node4	1.000	4	If f3d in [-1.323, -0.975[then Responders = 1 in 80% of cases
Node5	1.000	6	If f3d in [-0.975, -0.364[then Responders = 1 in 60% of cases
Node6	0.000	7	If f3d in [-0.364, 0.503[then Responders = 0 in 53.8% of cases
Node7	1.000	3	If f3d in [0.503, 1.105[then Responders = 1 in 60% of cases
Node8	0.000	2	If f3d in [1.105, 1.793[then Responders = 0 in 50% of cases
Node9	0.000	1	If czd in [-0.555, -0.34[and f3d in [-1.323, -0.975[then Responders = 0 in 100% of cases
Node10	1.000	4	If czd in [-0.34, 2.257[and f3d in [-1.323, -0.975[then Responders = 1 in 100% of cases
Node11	0.000	3	If c4d in [-1.406, -0.161[and f3d in [-0.975, -0.364[then Responders = 0 in 100% of cases
Node12	1.000	3	If c4d in [-0.161, 0.473[and f3d in [-0.975, -0.364[then Responders = 1 in 75% of cases
Node13	1.000	3	If c4d in [0.473, 0.728[and f3d in [-0.975, -0.364[then Responders = 1 in 100% of cases
Node14	1.000	3	If f8d in [-1.509, 0.519[and c4d in [-0.161, 0.473[and f3d in [-0.975, -0.364[then Responders = 1 in 100% of cases
Node15	0.000	1	If f8d in [0.519, 1.151[and c4d in [-0.161, 0.473[and f3d in [-0.975, -0.364[then Responders = 0 in 100% of cases

Node16	0.000	2	If fzd in [-1.985, -1.025[and f3d in [-0.364, 0.503[then Responders = 0 in 66.7% of cases
Node17	1.000	3	If fzd in [-1.025, 0.162[and f3d in [-0.364, 0.503[then Responders = 1 in 100% of cases
Node18	1.000	2	If fzd in [0.162, 0.715[and f3d in [-0.364, 0.503[then Responders = 1 in 66.7% of cases
Node19	0.000	4	If fzd in [0.715, 2.383[and f3d in [-0.364, 0.503[then Responders = 0 in 100% of cases
Node20	1.000	1	If f8d in [-2.184, -1.184[and fzd in [-1.985, -1.025[and f3d in [-0.364, 0.503[then Responders = 1 in 100% of cases
Node21	0.000	2	If f8d in [-1.184, 0.819[and fzd in [-1.985, -1.025[and f3d in [-0.364, 0.503[then Responders = 0 in 100% of cases

Tableau 9 – Règles de classification de la troisième analyse

Même si le support des règles est faible, certaines règles de classification semblent crédibles. Puisque cette analyse n'a couvert que dix-huit électrodes sur trente-cinq, il est difficile de déterminer la similarité de nos résultats avec ceux cités dans [14]. Cependant, en analysant l'arbre de décision (trop volumineux pour être inclus dans le document), on constate que la différence de valeurs de l'électrode F3 (qui se trouve dans la région préfrontale) entre la première et la deuxième session a été utilisée comme critère de fractionnement. Ce résultat est similaire à celui cité dans le travail de Rabinoff et al. [14] dans lequel il a été noté que la différence de valeurs de l'électrode FP1 (qui se trouve dans la région préfrontale) a été utilisée comme critère de fractionnement. Ce résultat confirme donc le rôle important des données EEG de la partie préfrontale dans la prédiction de la réponse aux traitements antidépresseurs. Puisque l'électrode F3 agit comme critère de fractionnement, il apparaît dans toutes les règles de classifications générées. De plus, sur les dix-huit électrodes utilisées dans cette analyse, seulement cinq (F3, CZ, C4, F8, FZ) intervenaient dans la génération des règles. Ces électrodes se trouvent respectivement dans les régions frontales *left frontal*, *right frontal*, *midline central temporal et midline frontal* (cf. la figure 3). Aucune électrode des régions postérieures n'a été choisie par l'arbre de décision. Cela confirme encore une fois l'importance des régions frontales dans la prédiction de la réponse aux traitements antidépresseurs.

5.2 Règles d'association avec négation et généralisation de variables

Nous avons aussi adapté des procédures de fouille de données développées au laboratoire LARIM de l'UQO sur les valeurs moyennes de l'EEG absolue des électrodes de chaque région cérébrale afin de générer des règles d'association avec négation sur des variables préalablement discrétisées puis généralisées dans l'espoir de découvrir de nouvelles associations et règles de classification. La capture d'écran suivante présente les règles les plus intéressantes obtenues (en gras).

Nous avons alors considéré les données thêta des 39 patients pour plusieurs régions cérébrales illustrées par la figure 3. Comme les données sont continues, une discrétisation a été produite sur les sept variables ci-dessous indiquées. Cela a généré plusieurs modalités (et intervalle de valeurs) pour chacune des variables. Nous avons ensuite procédé à une fusion d'intervalles pour obtenir un nombre réduit de modalités (2 à 3) représentant ainsi des attributs généralisés. Le tableau ci-après résume les variables considérées ainsi que le résultat de leur transformation. À titre d'exemple, la discrétisation de la variable RF a donné lieu à dix intervalles de valeurs qui ont été ensuite agrégés pour donner lieu à RFc1, RFc2 et RFc3 représentant respectivement les valeurs : faibles, moyennes et élevées des données thêta prélevées de la région frontale droite.

Codification Variables	Signification	c1	c2	c3
<i>RF</i>	<i>Right Frontal</i>	RFc1 : [0.915 - 1.319[RFc2 : [1.319- 1.92[RFc3 : [1.92 - 3.72[
<i>LF</i>	<i>Left Frontal</i>	[0.950 - 1.584[[1.584 - [1.835	X
<i>MF</i>	<i>Midline Frontal</i>	[0.226 - 1.436[[1.436 - 2.329[[2.329 - 4.280[
<i>LCT</i>	<i>Left Central/Temporal</i>	[0.784 - 1.070[[1.070 - 1.559[[1.559 - 2.041[
<i>LP</i>	<i>Left Posterior</i>	[0.660 - 1.449[[1.449 - 1.810[[1.810 - 2.148[
<i>RP</i>	<i>Right Posterior</i>	[0.551 - 1.967[[1.967 - 2.907[X
<i>MP</i>	<i>Middle Posterior</i>	[0.841 - 1.635[[1.635 - 2.219[X

Tableau 10 – Signification des variables

Dans un premier temps, nous avons généré l'ensemble des implications (c.-à-d. des règles d'association avec une confiance de 100%) avec un support d'au moins 20% (environ huit patients) et avons obtenu le tableau 11 ci-après dans lequel les règles de classification sont accentuées.

À titre d'exemple, la règle suivante signifie que si RPC2 est vrai (c.-à-d. RP a une valeur dans l'intervalle [0.551-1.967[), alors le patient est répondeur et a la propriété RFc3 (c.-à-d RF a une valeur au moins égale à 1.92). Le support de la règle est 20.51% et sa confiance est de 100%.

{R_{Pc2}} → {RESP, R_{Fc3}} 20.51% 100.0%

Association Rule Viewer

Context : ThetaGeneralRed

Min. support : 20.0%

Min. confidence : 100.0%

Rule count : 17

#	Antecedent	=>	Consequence	Support	Confidence
1.	{M _{Pc2} , RESP}	=>	{R _{Fc3} }	20.51%	100.0%
2.	{M _{Pc2} , R _{Fc3} }	=>	{RESP}	20.51%	100.0%
3.	{L _{Fc1} , R _{Fc3} }	=>	{RESP}	20.51%	100.0%
4.	{L _{Fc1} , M _{Fc3} }	=>	{RESP}	20.51%	100.0%
5.	{L _{Cc3} , M _{Fc3} }	=>	{R _{Fc3} }	20.51%	100.0%
6.	{R _{Pc2} }	=>	{RESP, R _{Fc3} }	23.07%	100.0%
7.	{L _{Cc3} , RESP}	=>	{R _{Fc3} }	23.07%	100.0%
8.	{R _{Fc1} }	=>	{M _{Pc1} , R _{Pc1} }	23.07%	100.0%
9.	{L _{Cc1} }	=>	{L _{Fc1} , L _{Pc1} }	25.64%	100.0%
10.	{M _{Fc3} , R _{Fc3} }	=>	{RESP}	28.2%	100.0%
11.	{L _{Pc1} }	=>	{M _{Pc1} , R _{Pc1} }	28.2%	100.0%
12.	{M _{Pc1} , R _{Fc2} }	=>	{L _{Fc1} }	33.33%	100.0%
13.	{R _{Fc2} }	=>	{R _{Pc1} }	35.89%	100.0%
14.	{L _{Fc1} , R _{Pc1} }	=>	{M _{Pc1} }	35.89%	100.0%
15.	{M _{Fc1} , M _{Pc1} }	=>	{L _{Fc1} }	38.46%	100.0%
16.	{L _{Cc2} , R _{Pc1} }	=>	{M _{Pc1} }	41.02%	100.0%
17.	{M _{Fc1} }	=>	{R _{Pc1} }	43.58%	100.0%

Export to XML...

Tableau 11 – Règles d’association y compris les règles de classification

Nous avons ensuite intégré la négation dans nos données de sorte que lorsque par exemple le patient n’a pas la propriété R_{Fc1}, nous ajoutons sa négation R_{Fc1}’ pour indiquer son absence. Cela nous permet d’établir des associations impliquant à la fois la présence et l’absence d’attributs ou propriétés.

Association Rule Viewer

Context : ThetaGeneralRedPN

Min. support : 20.0%

Min. confidence : 100.0%

Rule count : 96

#	Antecedent	=>	▽ Consequence	Support	Confidence
55.	{LFc2', MPc2'}	=>	{MPc1}	30.76%	100.0%
56.	{LPc3', RFc3'}	=>	{MPc2'}	51.28%	100.0%
57.	{LPc2', RFc2', RPe2'}	=>	{MPc2'}	20.51%	100.0%
58.	{LCc3', MFc2'}	=>	{MPc2'}	33.33%	100.0%
59.	{RESP', RFc2'}	=>	{MPc2'}	25.64%	100.0%
60.	{LFc2', LPc3', MFc3}	=>	{MPc2'}	20.51%	100.0%
61.	{LCc3', RPe2'}	=>	{MPc2'}	30.76%	100.0%
62.	{LPc2', LPc3'}	=>	{MPc2'}	33.33%	100.0%
63.	{LFc2', RPe2'}	=>	{MPc2'}	25.64%	100.0%
64.	{MPc1', RESP}	=>	{MPc2}	20.51%	100.0%
65.	{MFc3', RFc2', RPe2'}	=>	{RESP}	28.2%	100.0%
66.	{LCc2', MFc2', MFc3', RPe2'}	=>	{RESP}	20.51%	100.0%
67.	{RFc2', RFc3'}	=>	{RESP}	25.64%	100.0%
68.	{LCc1', MFc1', RFc1'}	=>	{RESP}	30.76%	100.0%
69.	{LPc1', LPc3', RFc1'}	=>	{RESP}	20.51%	100.0%
70.	{LFc2', LPc1', RFc1'}	=>	{RESP}	20.51%	100.0%
71.	{LFc2', RFc1', RFc2'}	=>	{RESP}	20.51%	100.0%
72.	{MFc1', MFc2', RFc1'}	=>	{RESP}	25.64%	100.0%
73.	{LPc3', RFc1', RFc2'}	=>	{RESP}	20.51%	100.0%
74.	{MPc2', RFc3}	=>	{RESP}	20.51%	100.0%
75.	{RFc3}	=>	{RFc1', RFc2'}	25.64%	100.0%
76.	{RFc2}	=>	{RFc1', RFc3'}	23.07%	100.0%
77.	{LCc2', MFc1'}	=>	{RFc1'}	23.07%	100.0%
78.	{LPc1', MFc1'}	=>	{RFc1'}	25.64%	100.0%
79.	{RESP}	=>	{RFc1'}	25.64%	100.0%

Export to XML...

Tableau 12 – Règles d’association avec négation y compris les règles de classification

À titre d’exemple, la règle de classification numéro 67 du tableau 12 indique que si le patient a une valeur de RF différente de RFc2 et RFc3, c.-à-d. une valeur de RF inférieure à 1.319, alors il est non répondeur avec un support de 25.64%, soit 10 sur 39 patients.

Si maintenant nous nous intéressons aux règles de discrimination, c.-à-d. celles qui ont comme prémisse la valeur de l’attribut de classification RESP (répondeur) ou sa négation RESP’ (non répondeur), nous nous concentrons sur la partie accentuée du tableau suivant.

Association Rule Viewer

Context : ThetaGeneralRedPN

Min. support : 10.0%

Min. confidence : 100.0%

Rule count : 120

#	∇ Antecedent	=>	Consequence	∇ Support	Confidence
87.	{MFc3', MPc1}	=>	{RFc3'}	41.02%	100.0%
88.	{MFc3', MPc2'}	=>	{RPc2'}	25.64%	100.0%
89.	{MFc3', RESP}	=>	{LPc3'}	10.25%	100.0%
90.	{MFc3', RFc2', RPc2'}	=>	{RESP}	12.82%	100.0%
91.	{MFc3}	=>	{MFc1', MFc2'}	25.64%	100.0%
92.	{MPc1', RESP}	=>	{MPc2}	20.51%	100.0%
93.	{MPc1'}	=>	{LPc1', RFc1'}	10.25%	100.0%
94.	{MPc1, RESP', RFc2'}	=>	{RFc1}	25.64%	100.0%
95.	{MPc1}	=>	{LPc3', MPc2'}	23.07%	100.0%
96.	{MPc2', RESP}	=>	{MPc1}	20.51%	100.0%
97.	{MPc2', RFc1', RFc2', RPc2'}	=>	{LPc2'}	10.25%	100.0%
98.	{MPc2', RFc1', RFc2'}	=>	{MFc2'}	15.38%	100.0%
99.	{MPc2', RFc3}	=>	{LCc3'}	38.46%	100.0%
100.	{MPc2, RESP}	=>	{RFc3}	20.51%	100.0%
101.	{MPc2, RFc3}	=>	{RESP}	20.51%	100.0%
102.	{MPc2}	=>	{MPc1'}	23.07%	100.0%
103.	{RESP', RFc1}	=>	{LPc2', MFc3'}	17.94%	100.0%
104.	{RESP', RFc2'}	=>	{MPc2'}	10.25%	100.0%
105.	{RESP}	=>	{RPc2'}	23.07%	100.0%
106.	{RESP, RFc2'}	=>	{MFc1'}	20.51%	100.0%
107.	{RESP, RPc2'}	=>	{LPc3'}	17.94%	100.0%
108.	{RESP}	=>	{RFc1'}	25.64%	100.0%
109.	{RFc1', RFc2', RPc2'}	=>	{MFc2'}	12.82%	100.0%
110.	{RFc1', RFc2'}	=>	{LCc1', RFc3}	25.64%	100.0%
111.	{RFc1', RFc3}	=>	{RFc2}	23.07%	100.0%
112.	{RFc1}	=>	{RFc2', RFc3'}	25.64%	100.0%
113.	{RFc2', RFc3}	=>	{RESP}	25.64%	100.0%
114.	{RFc2}	=>	{RFc1', RFc3'}	23.07%	100.0%
115.	{RFc3}	=>	{RPc2'}	23.07%	100.0%
116.	{RFc3}	=>	{RFc1', RFc2'}	25.64%	100.0%

Export to XML...

Tableau 13 – Règles d’association avec négation y compris les règles de discrimination

Les règles de discrimination trouvées avec un support d’au moins 20% sont les suivantes :

{RESP} => {RFc1'} 25.64% 100.0%

Interprétation: Si un patient répond efficacement à un traitement, alors nécessairement RFc1 n’est pas vrai. En d’autres termes, si un patient répond efficacement à un traitement, alors la valeur de RF est au moins égale à 1.319. Le support de la règle vaut 25.64%.

{RESP'} => {RPc2'} 23.07% 100.0%

L'interprétation de la règle est la suivante : Si un patient ne répond pas efficacement à un traitement, alors nécessairement RPc2 n'est pas vrai, c.-à-d. la valeur de RP est inférieure à 1.967. Le support de la règle vaut 23.07%.

Chapitre 6 Conclusion

Dans le cadre de ce mémoire, nous avons utilisé différentes techniques de fouille de données sur des électroencéphalogrammes de patients atteints de désordres dépressifs majeurs afin de découvrir des connaissances potentiellement inconnues par les cliniciens. Ces connaissances pourront ensuite être utilisées pour la prédiction précoce de la réponse des patients aux traitements et comme outil de diagnostic de la dépression. Une validation de nos résultats par les cliniciens a permis d'identifier les connaissances les plus pertinentes sous forme de règles de classification et d'association et sous forme de groupes (*clusters*) en matière de prédiction et de diagnostic.

À la demande de l'équipe du laboratoire LNRC et sur la base de notre étude de plusieurs articles scientifiques récents dans le domaine, nous avons mené de nombreuses analyses avec les données d'électroencéphalogrammes en valeurs absolues, relatives et selon la mesure de « cordance » qui combinent les deux formes. Nous avons aussi combiné les données sur les EEG avec les données cliniques (HAM-D et MADRS) pour obtenir de meilleurs résultats. À titre d'exemple, au lieu d'utiliser la bande de fréquence alpha seulement, nous avons également fait usage d'une combinaison de deux types de bandes de fréquences parmi alpha, bêta, thêta et delta et introduit davantage les valeurs de HAM-D et MADRS.

Après quelques transformations de données EEG, nous avons appliqué trois techniques de fouille de données : la classification, le regroupement et les règles d'association, et nous avons obtenu des résultats intéressants sur la prédiction précoce de l'efficacité d'un traitement antidépresseur. À titre d'exemple, nous avons pu confirmer les résultats cités dans des travaux précédents en identifiant la bande de fréquence thêta comme étant la bande la plus importante pour faire une classification d'un patient dans l'un des deux groupes : patients répondeurs et non répondeurs aux traitements. En particulier, nous avons constaté que lors de l'analyse des valeurs EEG de la bande thêta en combinaison avec les variables cliniques (HAM-D, MADRS, H17ChngW1B%, H17ChngW12%, MadChngW1B% MadChngW12B%), ces dernières n'ont pas apparu dans les règles de classification obtenues comme dans celles obtenues avec toutes les autres bandes. Cela indique l'importance de cette bande dans la prédiction de la réponse au traitement antidépresseur. De plus, les règles de classification et surtout les implications (avec une confiance de 100%) obtenues avec la bande thêta étaient celles qui avaient le plus haut support,

contrairement à celles obtenues avec toutes les autres bandes de fréquences. Nous avons aussi trouvé que la bande de fréquence alpha n'est pas utile pour la prédiction précoce de la réponse au traitement et que les règles de classification obtenues avec cette dernière avaient un bas support.

En outre, nos résultats ont confirmé l'importance de la région préfrontale du crâne dans la prédiction de la réponse des patients aux traitements antidépresseurs. Pour la bande de fréquence alpha total (8-14 Hz), c'est surtout l'électrode F4 (cf. figure 2) qui joue un rôle important dans la classification selon les deux groupes de répondeurs et non répondeurs.

Nous avons également introduit la négation et la généralisation au niveau des variables analysées et obtenu quelques nouvelles règles de classification et discrimination.

Finalement, les résultats obtenus au niveau du diagnostic de la maladie ont été moins concluants malgré la prise en compte de diverses techniques et combinaisons de variables.

RÉFÉRENCES

- [1] Dan Vlad Iosifescu, Electroencephalography – Derived Biomarkers of Antidepressant Response, *Harvard Review of Psychiatry*, Vol. 19, No. 3, Février 2011.
- [2] Pang-Nin Tan, Michael Steinbach, Vipin Kumar, Introduction to Data mining, Première édition, Addison Wesley, Boston, 2006.
- [3] Jiawei Han, Micheline Kamber, Jian Pei, Data mining Concepts and Techniques, troisième édition, Morgan Kaufmann, Waltham, 2012.
- [4] Bernhard Ganter, Rudolf Wille, Formal concept analysis - mathematical foundations, Première édition, Springer, 1999.
- [5] Glenn J. Myatt, Making sense of data, Première édition, John Wiley & Sons, New Jersey, 2007.
- [6] Knott VJ, Telner JI, Lapierre YD, Browne M, Horn ER, Quantitative EEG in the prediction of antidepressant response to imipramine, *J Affect Disord* 1996;39:175–84.
- [7] Bruder GE, Stewart JW, Tenke CE, et al. Electroencephalographic and perceptual asymmetry differences between responders and non responders to an SSRI antidepressant. *Biol Psychiatry* 2001; 49:416–25.
- [8] Bares M, Brunovsky M, Kopecek M, Novak T, Stopkova P, Novak T, Kozeny J, Sos P, Krajca V, Hoschl C. Early reduction in prefrontal theta QEEG cordance value predicts response to venflaxine treatment in patients with resistant depressive disorder. *European Psychiatry* 2008; 23:350-355. [PubMed: 18450430].
- [9] Bares M, Brunovsky M, Kopecek M, Novak T, Stopkova P, Novak T, Kozeny J, Sos P, Krajca V, Hoschl C. Changes in QEEG prefrontal cordance as a predictor of response to antidepressants in patients with treatment resistant depressive disorder: A pilot study. *Journal of Psychiatric Research* 2007; 41:319-335. [PubMed: 16889798].
- [10] Ian A. Cook et al. Midline and right frontal brain function as a physiological biomarker of remission in major depression. *Psychiatry Res.* 2009 November 30; 17(2):152-157.
- [11] Knott VJ, Lapierre YD. Computerized EEG correlates of depression and antidepressant treatment. *Progress in Neuropsychopharmacology and Biological Psychiatry* 1987; 11:213-221.
- [12] Knott VJ, Telner JI, Lapierre YD, Browne M, Horn ER. Quantitative EEG in the prediction of prediction of antidepressant response to imipramine. *Journal of Affective Disorders* 1996; 39:175-184. [PubMed: 8856421].

[13] Pizzagali D et al. Anterior cingulate activity as a predictor of degree of treatment response in major depression: Evidence from brain electrical tomography analysis. *American Journal of Psychiatry* 2001; 158:405-415. [PubMed: 15288493].

[14] Rabinoff M, Kitchen C.M.R, Cook I.A. and Leuchter A.F. Evaluation of Quantitative EEG by Classification and Regression Trees to Characterize Responders to Antidepressant and Placebo Treatment. *The Open Medical Informatics Journal* 2011; 5: 1-8.

[15] Cook IA, Leuchter AF, Morgan ML, Stubbeman W, Siegman B, Abrams M. Changes in prefrontal activity characterize clinical response in SSRI nonresponders: a pilot study. *J Psychiatr Res.* 2005; 39: 461-6.

[16] Cook IA, Leuchter AF, Morgan M, *et al.* Early changes in prefrontal activity characterize clinical responders to antidepressants. *Neuropsychopharmacology* 2002; 27: 120-31.