
UNIVERSITÉ PARIS 8 - VINCENNES À SAINT-DENIS

Licence Informatique Parcours MIME

RAPPORT DE STAGE

Conception d'un système embarqué pour la classification des signaux EEG

Par : Youcef MCIRDI.

Encadrant : : Larbi BOUBCHIR
Organisme d'accueil : Laboratoire d'informatique avancée de Saint-Denis (LIASD)

Table des matières

1	Remerciments	1
	Introduction	1
2	Les Signaux EEG	3
2.1	Qu'est ce que EEG ?	4
2.2	Les techniques d'acquisition des signaux EEG	4
2.2.1	Le système de localisation des électrodes (Système normalisé 10-20)	5
2.2.2	Les différentes catégories d'acquisition	5
	Le Système d'acquisition direct/indirect	5
	Le système d'acquisition Synchrone/Asynchrone	7
2.3	EEG normal	8
2.4	EEG anormal et épileptique	9
2.5	Domaine d'application EEG	10
2.5.1	Neuromarketing :	10
2.5.2	Facteurs humains :	11
2.5.3	Interaction sociale :	11
2.5.4	Psychologie et neurosciences :	11
2.5.5	Études cliniques et psychiatriques :	12
2.5.6	BCI ou ICM	12
2.6	Conclusion	12
3	Conception et conclusion	15
3.1	Introduction	15
3.2	Matériel et Données	15
3.3	Prétraitement	16
3.4	Classification	17
3.4.1	Formules utilisées	17
3.4.1.1	La somme	17
3.4.1.2	La moyenne arithmétique	17

3.4.1.3	La variance	17
3.4.1.4	L'écart-type	18
3.4.1.5	La médiane	18
3.4.1.6	La médiane haute	18
3.4.1.7	La médiane basse	18
3.4.1.8	Le taux de passage a zéro	18
3.4.1.9	L'asymétrie (skewness)	19
3.5	Réseaux neurones pour la classification	19
3.6	Resultat et Conclusion	21
4	Quelques extraits de code	25
5	Référnces	27

Table des figures

3.1	K-Nearest Neighbors	21
-----	-------------------------------	----

Chapitre 1

Remerciements

Je tiens à remercier dans un premier temps, toute l'équipe pédagogique de l'université Paris-8 et les intervenants professionnels responsables de la formation Informatique, pour avoir assuré la partie théorique de celle-ci.

Je remercie également Madame SEDDIKI Lynda et Monsieur TOUATI Youcef pour l'aide et les conseils concernant les missions évoquées dans ce rapport, qu'ils m'ont apporté lors des différents suivis.

Je tiens à remercier tout particulièrement et à témoigner toute ma reconnaissance aux personnes suivantes, pour l'expérience enrichissante et pleine d'intérêt qu'elles m'ont fait vivre durant ces deux mois au sein de laboratoire d'informatique avancée de saint denis.

Monsieur Boubchir Larbi, Maitre de conference et mon tuteur de stage, pour son accueil et la confiance qu'il m'a accordé, pour le temps qu'il m'a consacré tout au long de cette période, sachant répondre à toutes mes interrogations ; sans oublier sa participation au cheminement de ce rapport.

Je tiens à remercier madame BENSLIMANE Itaf pour son accompagnement durant mon stage.

Chapitre 2

Introduction

Passionné par tout ce qui est Interface Homme-Machine ,Big Data et Intelligence artificielle j'ai effectué mon stage de fin d'études au LIASD au sein l'université Paris-8, mon tuteur de stage monsieur Boubchir m'a proposé de faire une recherche sur les signaux EEG et de réaliser un système embarqué de classification des signaux EEG pour la detection des crises épileptique. L'épilepsie est l'une des maladies neurologiques les plus fréquentes, touchant environ 1% de la population (on estime que 600 000 personnes souffrent d'épilepsie en France). Dans l'état actuel des la connaissances médicales, près de 30% des patients épileptiques ne peuvent être traités par des médicaments ou chirurgie.

Ils peuvent souffrir d'une crise soudaine, n'importe quand, n'importe où une crise peut durer quelques secondes ou quelques minutes, et a une forte incidence sur le fonctionnement de leur système nerveux, la mémoire et la conscience ; en conséquence, les aspects sociaux et les capacités professionnelles des patients et de leurs familles diminuaient fortement.

Un électroencéphalogramme est un ensemble de signaux électriques (figure 1), à l'échelle de microvolts, prélevés par des électrodes insérées dans le cerveau par chirurgie (Fig. 2b) ou collées. au cuir chevelu (Fig. 2a).

Il est supposé que de nombreuses informations neurologiques sont intégrées dans les signaux EEG. Pour prédire ou détecter une crise, il est nécessaire de développer des méthodes d'analyse de signaux capables de détecter des modèles dans l'EEG typiques d'un état pré-ictal ou d'un état ictal, respectivement.

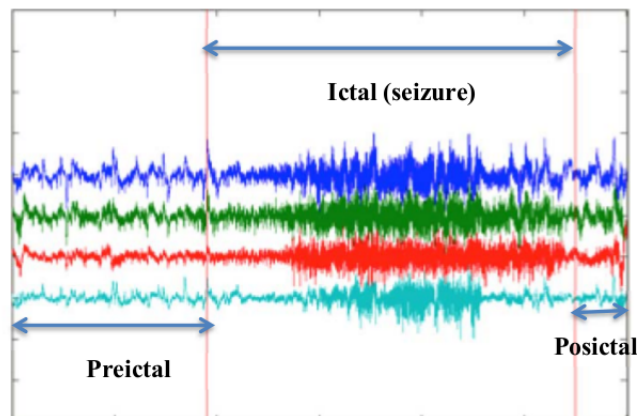
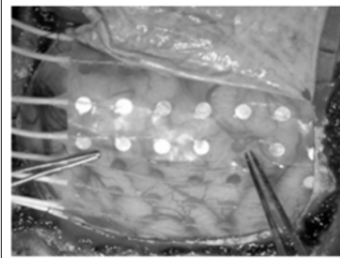


Figure 1. EEG avec une crise



(a)



(b)

Figure 2. (a) Électrodes de surface (non invasives), (b) Électrodes intra-crâniennes (invasives).
Chaque électrode alimente un canal EEG par rapport à une référence électrique.

Chapitre 3

Les Signaux EEG

Le cerveau humain est toujours actif. Le cerveau contrôle les différentes activités du corps. Les fonctions cérébrales peuvent être surveillées en observant les signaux électriques générés dans les neurones. Ce signal est appelé électroencéphalogramme (EEG). Le signal peut être extrait à l'aide d'électrodes et peut être visualisé à l'aide d'un voltmètre, d'un oscilloscope ou sur un écran d'ordinateur. Le signal EEG peut être enregistré et le phénomène est connu sous le nom d'électroencéphalographie. Le signal EEG peut être utilisé pour étudier l'état du cerveau humain et la santé globale de la personne.

Ces signaux sont approximativement inférieurs à 100 V et 100 Hz et peuvent être mesurés avec des électrodes placées sur le cuir chevelu, de manière non invasive. En raison de leur faible amplitude due à la composition du crâne, la mesure de l'EEG est plus difficile que les autres mesures non invasives de signaux biologiques comme l'électrocardiogramme (ECG), l'électromyogramme (EMG), l'électrooculogramme (EOG) etc. Le fait d'avoir des systèmes d'enregistrement de signaux biologiques coûteux ne peut pas garantir l'acquisition de signaux appropriés.

L'EEG a été inventé dans les années 1920 par un psychiatre allemand, le Dr Berger, mais ne s'est vraiment développé et répandu dans la pratique médicale courante, en particulier dans la pratique épiléptologique, que dans les années 1950. Le principe de l'EEG est de recueillir les potentiels électriques sur un appareil qui amplifie les signaux, puis les transcrit pour qu'ils puissent être analysés.

3.1 Qu'est ce que EEG ?

L'électroencéphalographie (EEG) est une technique qui consiste à enregistrer l'activité électrique d'une partie des neurones du cortex. Un bonnet, qui comporte quelques dizaines d'électrodes, est placé sur la tête du patient et relié à un système d'acquisition. À chaque électrode correspond un signal qui représente, en temps réel, l'activité moyenne des neurones présents à quelques centimètres sous cette électrode. Cette méthode permet de visualiser l'activité du cerveau ; elle est utilisée pour suivre le déclenchement et la propagation des crises, apportant des informations complémentaires à celles provenant de la sémiologie.

À ses débuts, l'EEG étant directement tracé sur des bandes de papier à la manière d'un sismographe. Avec le développement des systèmes numériques à partir des années 90, on peut maintenant facilement stocker, et analyser les signaux EEG, de manière visuelle ou automatique. Les technologies actuelles permettent aussi d'augmenter la résolution spatiale en plaçant jusqu'à 256 électrodes sur le scalp, et la résolution temporelle, avec des fréquences d'échantillonnage qui montent souvent à 2 kHz

L'EEG est très utile, mais ne permet pas de visualiser l'activité de neurones situés plus en profondeur dans le cerveau. Pour pallier ce problème, Jean Bancaud et Jean Talairach ont élaboré, en 1962, une technique chirurgicale qui consiste à implanter des électrodes à l'intérieur du crâne, directement dans le cerveau (Talairach et al. 1962). Cette méthode appelée Stéréo-électroencéphalographie (SEEG) utilise le repérage dit stéréotaxique pour placer les électrodes précisément dans les régions souhaitées. Cette méthode permet d'augmenter la résolution spatiale de l'EEG de scalp en mesurant l'activité de neurones inaccessibles auparavant ; ce qui la rend très pratique pour certaines épilepsies, notamment pour déterminer la zone de départ des crises lorsqu'elles impliquent des régions sous-corticales. Cependant, et contrairement à l'EEG de scalp, la SEEG est invasive ; c'est pourquoi elle n'est utilisée qu'en prévision d'une chirurgie. La SEEG est effectuée pour identifier précisément les zones pathologiques et pour faire une cartographie fonctionnelle de cette zone. Cette dernière étape permet ensuite d'estimer les chances de réussite et les risques, afin de proposer une éventuelle chirurgie si ces risques sont jugés raisonnables et inférieurs aux bénéfices éventuels.

3.2 Les techniques d'acquisition des signaux EEG

L'acquisition des signaux EEG représente une étape très importante qui précède l'analyse et le traitement. En effet, l'interprétation précise de ces

3.2. LES TECHNIQUES D'ACQUISITION DES SIGNAUX EEG

signaux n'est pas possible en absence d'une connaissance approfondie des emplacements des électrodes sur le cuir chevelu. Les électrodes doivent être placées dans les régions où il y a une forte activité neuronale.

3.2.1 Le système de localisation des électrodes (Système normalisé 10-20)

La localisation des électrodes peut se faire manuellement ou en utilisant des appareils de placement dédiés. En effet, la méthode manuelle consiste à identifier manuellement la position des électrodes puis à calculer leurs coordonnées en 3D. Alors que la deuxième méthode calcule la position 3D des électrodes soit en utilisant des ondes électromagnétiques, acoustiques, ou à partir d'une analyse de la morphologie de la tête du patient. La localisation doit être précise, reproductible et peu onéreuse. Pour cela, la méthode manuelle est, souvent, la plus utilisée. Cette technique suppose que les électrodes EEG sont positionnées dans une configuration définie selon les systèmes internationaux 10-20 ou 10-10. Le système 10-20 est un système de localisation international proposé comme standard par la société Américaine de l'électroencéphalographie. Le système décrit le placement des électrodes selon des positions bien définies le long de la surface de la tête. La figure 1.2a montre le positionnement des électrodes du système 10-20 où la surface de la tête est divisée en plusieurs régions. Chaque position, liée à une électrode sur le crâne, est déterminée par une lettre majuscule, pour identifier la région, suivie d'un nombre ou d'une petite lettre pour identifier l'emplacement de l'hémi-sphère. Les lettres F, T, C, P et O se réfèrent respectivement aux termes : Frontal, Temporal, Central, Pariétal et Occipital. Les numéros pairs réfèrent l'hémisphère droit (resp gauche). Z indique que l'électrode est placée sur la ligne médiane. Les numéros 10 et 20 représentent les pourcentages des distances qui séparent deux électrodes adjacentes par rapport à la distance totale de l'inion (arrière) à la Nasion (devant) du crâne.

3.2.2 Les différentes catégories d'acquisition

Le Système d'acquisition direct/indirect

Le Système d'acquisition direct/indirect La technique d'acquisition dite directe consiste à implanter les capteurs au niveau de la surface de contact avec le cerveau par le biais d'une intervention chirurgicale assez délicate. Cette dernière, connue par la craniotomie, requiert l'ouverture du crâne ainsi que l'incision des différentes couches des membranes qui recouvrent le cerveau

. Une fois les électrodes sont mises en contact direct avec les neurones, elles délivrent l'activité électrique post-synaptique et/ou les potentiels d'action émis sur l'axone du neurone . Bien qu'elle est assez délicate, cette technique est caractérisée par une excellente qualité du signal, une très bonne résolution temporelle et une bande de fréquence assez large. De plus, la présence des bruits et des artefacts est réduite à son minimum. Le recours à cette technique d'acquisition est dictée par le besoin d'obtenir des mesures de potentiels cérébrales assez précises pour des applications de diagnostic des pathologies médicales comme l'Alzheimer, l'épilepsie et les troubles neuronales. Quant à la méthode d'acquisition indirecte, connue aussi par la technique d'acquisition non-invasive, elle se contente de placer les capteurs sur le cuir chevelu pour enregistrer l'activité cérébrale produite suite à une imagination d'une tâche ou une exécution d'un mouvement donné. Cette méthode demeure la plus sollicitée par les systèmes d'acquisition depuis plusieurs décennies vu sa facilité d'application, son faible coût ainsi que l'absence d'intervention chirurgicale. De plus, elle est dotée d'une bonne résolution temporelle permettant de détecter le moindre changement au niveau des rythmes cérébraux. En revanche, la résolution spatiale ainsi que la gamme de fréquences obtenues par cette technique est limitée puisqu'elle est vulnérable à la présence d'artefacts (bruits et interférences entre des signaux provenant d'autres activités cérébrales non voulues). Pour augmenter la sensibilité de ces électrodes et améliorer la qualité des signaux EEG, un gel est appliqué entre les électrodes et le cuir chevelu (électrodes mouillées). De plus, l'application des techniques de filtrage et d'amplification ont rendu cette technique mieux adaptée pour un spectre assez large d'applications telles que : les ICMS, les applications de communication ainsi que les applications de diagnostic des pathologies médicales. La figure suivante présente une illustration de deux méthodes d'enregistrement directe et indirecte de l'activité électrique à partir des zones actives du cerveau.

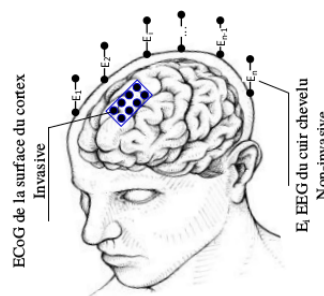


Figure -Les techniques d'acquisition directes et indirectes des signaux EEG

3.2. LES TECHNIQUES D'ACQUISITION DES SIGNAUX EEG⁹

Le système d'acquisition Synchrone/Asynchrone

La qualification du système d'acquisition à base d'ICMs est établie selon le paradigme de communication entre l'utilisateur et la machine. Ainsi, un système est dit synchrone lorsqu'il oblige son utilisateur à communiquer avec la machine durant des intervalles de temps bien définis [GAP10]. Ces intervalles de temps sont indiqués à l'utilisateur par des signaux visuels ou acoustiques. La figure 1.4 schématise le diagramme de temps utilisé dans la compétition "BCI- competition" pendant l'acquisition des signaux EEG. Le processus d'enregistrement commence par un signal sonore, qui informe l'utilisateur du démarrage de la session d'enregistrement. D'autres systèmes d'acquisition utilisent le clavier visuel pour aider l'utilisateur à suivre la synchronisation des différentes étapes d'enregistrement. L'utilisateur en question doit se préparer aux étapes suivantes d'enregistrements effectives qui durent 7.5 secondes. Durant cette période (trial), il est censé effectuer une tâche mentale liée au mouvement de la main droite ou gauche. L'utilisation de cette technique d'acquisition nécessite une phase d'apprentissage par ses utilisateurs.

Par opposition au système synchrone, un système est dit asynchrone lorsque l'enregistrement de l'activité cérébrale pourrait survenir à tout moment sans se soucier de n'importe quel signal de synchronisation [Bek+09]. Dans ce cas l'utilisateur a le choix d'envoyer ou non un signal indiquant le début de l'action. Le système asynchrone est censé effectuer des analyses et des traitements des signaux EEG continuellement. Ce système requiert un apprentissage de longue durée pouvant s'étendre à plusieurs heures. À la suite du traitement des signaux EEG, le système visualise en temps réel les signaux acquis ainsi que les topographes représentant les caractéristiques de ces signaux. Ce mode d'acquisition est plus exigeant, mais il est plus naturel, confortable et beaucoup plus pratique. La commande d'un tel système asynchrone par les signaux sensorimoteurs liés au mouvement de l'un des organes humains nécessitent la localisation du rythme et . Pour cela le système analyse continuellement les signaux acquis afin d'identifier ces rythmes.

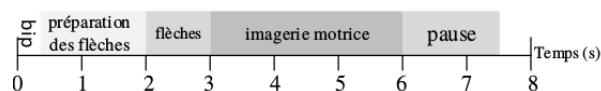


Figure -Flot d'enregistrement des signaux EEG dans le système synchrone

3.3 EEG normal

Les signaux ainsi obtenus présentent des oscillations plus ou moins régulières en fonction de la synchronisation des cellules corticales. Leurs amplitudes sont de l'ordre de 10 à 200 μV et leurs fréquences vont typiquement de 0.5 à 50 Hz. On distingue 4 catégories principales d'ondes EEG en fonction de leurs rythmes et amplitudes.

Rythme	Fréquence	Amplitude	Corrélat
Delta	< 4 Hz	>30 μV	Sommeil lent profond
Thêta	4-7 Hz	20 μV	Sommeil léger
Alpha	8-12 Hz	30 μV	Veille calme
Beta	15-30 Hz	20 μV	Veille active

Les différents rythmes physiologiques.

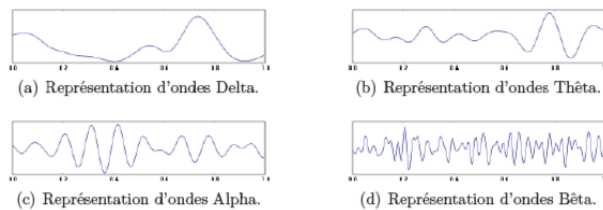
-Les ondes beta sont présentes lorsque l'on est éveillé et que l'on garde les yeux ouverts.

Elles apparaissent de manière symétrique sur la tête et leurs amplitudes sont faibles.

-Le rythme alpha est une activité normale d'un adulte éveillé, mais il est surtout présent lorsque l'individu se relaxe et garde les yeux fermes. Il apparaît simultanément des deux côtes de la tête généralement plus vers l'arrière qu'à l'avant, et est donc mieux visualisé sur un EEG occipital (Oz-A1) que sur un EEG central (Cz-A1) ou Frontal (FP1-A1). Il disparaît dès que le sujet ouvre les yeux ou en cas d'activité mentale

-L'activité thêta peut aussi bien être normale qu'anormale en fonction de l'âge et de la pathologie de la personne. En effet, elle est tout à fait naturelle s'il s'agit d'un adulte et que celui-ci est endormi. Par contre, si ce dernier est parfaitement éveillé, la présence d'ondes thêta peut être le signe d'un dysfonctionnement du cerveau.

- Le rythme delta n'est normal que si l'individu est plongé dans un sommeil profond. Les ondes correspondantes présentent alors des amplitudes généralement plus élevées que pour toutes les autres ondes et sont plus marquées sur un EEG frontal (FP1-A1) que central (Cz-A1)



Principaux rythmes oscillatoires de l'EEG humain

Les anomalies du tracé consistent en une modification des fréquences et de l'amplitude, permanente ou transitoire, du tracé, et en la présence de figures anormales.

3.4 EEG anormal et épileptique

EEG anormal des enfants prématurés Les anomalies de l'EEG prématuré peuvent être caractérisées comme des anomalies aiguës ou chroniques. Cette catégorisation de l'EEG s'est révélée très utile puisque des changements aigus de l'activité électro-corticale apparaissent après des sévères problèmes pathologiques. Les anomalies persistantes du stade chronique sont souvent des marqueurs de lésions établies . Les anomalies aiguës apparaissent pendant et après un changement dans la continuité (IBI plus long et absence des activités continus), ou une fréquence plus rapides et une amplitude réduite . La présence de telles caractéristiques d'EEG a été associée à des lésions cérébrales hémorragiques et ischémiques chez les nourrissons prématurés . Maruyama et al. Ont montré chez des nourrissons très prématurés que le degré maximal d'anomalies aiguës au cours des premiers jours postnataux était associé au risque de développement ultérieur de la paralysie cérébrale et était également corrélé avec sa sévérité . 24 Les anomalies chroniques indiquent des anomalies persistantes et caractéristiques de l'EEG apparaissant lorsque les anomalies aiguës ont disparu. Il s'agit notamment des activités désorganisées qui ont été associés à un mauvais développement neurologique chez les nourrissons très prématurés et modérés . L'augmentation de l'IBI seule est difficile à attribuer à des anomalies aiguës ou chroniques car elle peut représenter soit une suppression aiguë, soit une activité immature. Des activités désorganisées sont caractérisés par des ondes anormales, par exemple ondes rolandiques positives, qui sont prédictives si elles sont fréquentes suite à une paralysie cérébrale.

EEG du patient épileptique revêt trois états différents, L'état de crise est appelé « **phase critique** » , elle peut être précédée par une phase prodromique appelée « aura ». La phase critique est suivie d'une « **phase postcritique** » correspondant au retour du sujet vers un état normal qui peut durer de quelques minutes à plusieurs heures. La période qui sépare l'occurrence de deux crises est appelée « **phase intercritique** » .

L'électroencéphalogramme permet donc de détecter les perturbations d'origine fonctionnelle ou lésionnelle des aires corticales de la convexité, d'apprécier leur importance, de préciser éventuellement leur localisation et de suivre leur évolution. C'est aussi un examen performant pour étudier d'une part

les troubles de la vigilance, d'autre part les expressions paroxystiques plus ou moins rattachées aux manifestations épileptiques, mais il reste peu spécifique et ce n'est qu'exceptionnellement qu'il permet un diagnostic étiologique.

3.5 Domaine d'application EEG

Bien que la technologie de l'EEG existe depuis la presque fin du siècle dernier, ce n'est que très récemment qu'elle a été mise à la disposition du grand public.

Jusqu'à il y a quelques décennies, seuls les cliniciens et les chercheurs experts en neurosciences étaient capables de mettre en place et d'analyser des enregistrements dans des conditions de laboratoire strictement contrôlées avec des bouchons d'électrodes contenant 64 canaux ou plus. Ces électrodes de haute densité uniformément espacées sur le cuir chevelu humain ont été utilisées pour aider les chercheurs à découvrir les mécanismes neuronaux sous-jacents impliqués dans les actions, la cognition ou le traitement émotionnel étaient complètement inconnus.

Les progrès récents dans le domaine du matériel informatique et de la technologie des processeurs ont permis aux chercheurs du monde entier d'élargir considérablement les connaissances existantes sur la complexité du cerveau humain et de mieux comprendre les processus et les structures du cerveau.

Maintenant que cette pierre angulaire a été posée, l'EEG (électroencéphalographie) peut être utilisée pour diverses applications.

3.5.1 Neuromarketing :

Dans le domaine du neuromarketing, les économistes utilisent la recherche EEG pour détecter les processus cérébraux qui déterminent les décisions des consommateurs, les zones du cerveau qui sont actives lorsque nous achetons un produit/service, et les états mentaux dans lesquels se trouve la personne concernée lorsqu'elle explore des réserves physiques ou virtuelles. Aujourd'hui, des études peuvent être menées dans des installations mobiles pour obtenir des informations sur les habitudes d'achat et la prise de décision dans des scénarios du monde réel.

3.5.2 Facteurs humains :

Issu de la psychologie, le domaine des facteurs humains se concentre sur l'optimisation du milieu de travail, tant au niveau des outils et des interfaces qu'au niveau de l'interaction sociale. Dans ce domaine, la recherche en EEG est utilisée pour identifier les processus cérébraux liés à des traits de personnalité spécifiques tels que l'introversion/extroversion ou l'anxiété sociale.

De plus, les processus cérébraux reflétant les états cognitifs et attentionnels au cours de l'interaction homme-machine sont fortement étudiés à l'aide de l'EEG, principalement à l'aide de casques sans fil dotés de capacités de surveillance à long terme.

3.5.3 Interaction sociale :

Les humains sont des agents sociaux - nous passons la majorité de notre vie à interagir avec les autres. Dans la recherche sur les interactions sociales, les processus cérébraux liés à la perception sociale, à l'auto-évaluation et au comportement social sont étudiés. Il est important de noter que les interactions sociales et la communication ne sont pas des formes passives de traitement des stimuli entrants.

Chaque fois que nous parlons aux autres ou que nous résolvons des problèmes ensemble, nous devons nous "synchroniser" avec nos partenaires. Pour étudier les processus cérébraux qui sous-tendent la synchronisation des conversations et des actions, les chercheurs de l'EEG utilisent une méthode appelée "hyperbalayage" pour enregistrer les données de plusieurs personnes à la fois, ce qui leur permet de mieux comprendre les interactions entre le leadership et l'équipe.

3.5.4 Psychologie et neurosciences :

De façon générale, les études psychologiques utilisent l'EEG pour étudier les processus cérébraux qui sous-tendent l'attention, l'apprentissage et la mémoire. Comment percevons-nous le monde? Comment nos attentes façonnent-elles la façon dont nous voyons notre environnement?

Sur la base d'une répétition massive des essais, des potentiels liés à l'événement sont extraits du flux continu de données EEG, ce qui permet de caractériser les processus cérébraux déclenchés par les événements sur une échelle de temps très détaillée (des dizaines de millisecondes).

Les PRE peuvent être caractérisés par leur amplitude (en millivolts, avec des ondes positives et négatives étiquetées "P" et "N", respectivement), leur synchronisation (en ms par rapport au début de l'événement) et leur distribution de tension entre toutes les électrodes (topographie). Des ERP spécifiques ont été identifiés pour le traitement des visages (N170), des mots et significations (N400), de la surprise (P300) ou du rappel de mémoire (P600).

3.5.5 Études cliniques et psychiatriques :

Lorsque les processus cérébraux sont altérés (p. ex. lésions, dysfonctionnements génétiques, maladies), des déficits dans le traitement comportemental, attentionnel et cognitif peuvent être observés. Les domaines cliniques et psychiatriques utilisent l'EEG pour évaluer les états cognitifs des patients, déterminer les sites de lésions et classifier les symptômes.

De plus, l'EEG est largement utilisé pour évaluer l'effet du traitement médical et psychologique (p. ex. en thérapie cognitivo-comportementale). De plus en plus de thérapies utilisent la technologie de la réalité virtuelle et enregistrent les données EEG pour surveiller la façon dont le cerveau des patients s'améliore avec le temps.

3.5.6 BCI ou ICM

Les interfaces cerveau-ordinateur constituent un domaine relativement nouveau mais émergent pour l'EEG. Aujourd'hui, nous savons plus en détail quelles zones du cerveau sont actives lorsque nous percevons des stimuli, lorsque nous préparons et exécutons des mouvements corporels, ou lorsque nous apprenons et mémorisons des choses.

Cela donne lieu à des applications EEG très puissantes et ciblées pour piloter des appareils utilisant l'activité cérébrale. Cela peut, par exemple, aider les patients paralysés à diriger leur fauteuil roulant ou à déplacer un curseur sur un écran, mais la technologie BCI est également utilisée dans les scénarios militaires où les soldats sont équipés d'un exosquelette et d'une capsule EEG, leur permettant de déplacer, soulever et transporter des objets très lourds simplement en fonction de leur activité cérébrale.

3.6 Conclusion

Ce chapitre présente des notions générales sur les signaux EEG, la technique d'acquisition, domaine d'application et diagnostic principalement l'EEG,

la technique la plus couramment utilisée car elle permet de recueillir un grand nombre d'information au cours de la crise ou entre les crises (phase intercritique).

Chapitre 4

Conception et conclusion

4.1 Introduction

De nombreuses personnes dans le monde sont atteintes d'épilepsie, ce qui représente de la population totale. Une attaque peut se produire à tout moment, elle peut aller de quelques secondes à plusieurs minutes, affectant ainsi de nombreuses parties du cerveau. Aujourd'hui, avec l'aide de l'EEG, la possibilité de prédire ou détecter une crise d'épilepsie devient un défi pour la recherche. Si ça pouvait d'élaborer un algorithme permettant de détecter à temps une crise, de nouvelles stratégies de détection de la le désarmement de la saisie pourrait éventuellement être développé. De l'autre côté, la saisie la prédiction permettrait au patient de prendre des mesures pour sa propre sécurité et son bien-être social. pendant la crise. Comme point de départ ont été recueillis EEG des patients au moment de la crise précédente, pendant, après et entre les crises.

4.2 Matériel et Données

Les données recueillies étaient représentées sous formes de des fichiers TXT ;quelques un contiennent les données des cas normaux et les autres des cas anormaux, ou chaque ligne représente un signal. les colonnes des fichiers représentent l'amplitude du signal de chaque cas et chaque fichier TXT est constitué de plusieurs colonnes.

Les caractéristiques ont été extraites à l'aide de segments de 2 secondes d'EGE à 0se chevauchent. Concrètement, toutes les secondes, un vecteur avec 29 bandes de fréquences est généré. Pour cette raison, nous obtenons une matrice où les colonnes représentent les colonnes extraites. et les lignes un

vecteur de caractéristiques. Chaque valeur de ce vecteur représente différents états :

- 0, représentant un état non ictal(non cris). Cela signifie que contais inter-ictal, pré-ictal, pré-ictal et les états pos-ictal

- 1, ne représente que l'état ictal. Cela signifie qu'il y a une crise. Ces fichiers sont importés de la BD CHB-MIT sous forme .edf et convertis avec logiciel EDF BROWSER. Chacun de ces fichiers TXT contient une matrice, chaque ligne représente un signal numérique EGG d'une durée de 1 secondes. Ce sont des signaux de patients de différents âge et sexe venant de différents électrocardiogramme. la fréquence d'échantillonnage est de 125 HZ.

- Logiciel : EDF BROWSER
- Plateforme : MBED
- Langage de programmation : Matlab, C.
- Carte MBED.

4.3 Prétraitement

Pour prédire les crises, nous avons dû diviser la matrice cible, qui n'a plus qu'à deux états ("non-ictal" et "ictal"), en quatre états. Pour ce faire, nous considérons que 10 minutes (équivalent à 600 points) avant chaque crise en tant qu'état d'"état" pré-Ictal. où la crise est sur le point de se produire" et 5 minutes après la fin de chaque crise (équivalent à 300 points) en tant qu'état Post-Ictal dans lequel la crise a été vient d'arriver". Les points restants à 0 sont maintenant représentés en tant qu'état Inter-Ictal "état normal du patient" et les points à 1 restent à l'état Ictal "état dans lequel l'état du patient est le suivant le patient est en crise." Avec cette répartition, nous avons obtenu un nouvel objectif avec les colonnes possibles suivantes :

- - État inter-Ictal, (1,0,0,0,0)
- - État pré-ictal, (0,1,0,0,0)
- - État intestinal, (0,0,1,0)
- - État postinctal, (0,0,0,0,1)

Pour le CNN : une matrice 29x29 basée sur l'ensemble de données, avec une recible spéculative.

4.4 Classification

La classification a pour but d'identifier les classes auxquelles appartiennent des objets à partir des attributs. Il existe plusieurs types de classifications. Les plus connus sont : supervisée et non-supervisée. la classification supervisée est celle qui se base sur une référence (cas déjà traités et dont les résultats sont connus et répertoriés). Les cas sont associées à des labels (dans notre cas il s'agit de bon et de mauvais signaux).

L'objectif est de permettre au classifieur, en se basant sur un modèle spécifique et un ensemble des règles, de faire la différences entre les cas normaux et anormaux. sur cette base, le classifieur pourrait traiter et affecter les résultats de nouvelles données (nouveaux signaux) dans soit la classe normale ou anormale, en utilisant comme référence les cas déjà répertoriés.

4.4.1 Formules utilisées

4.4.1.1 La somme

Pour additionner tous les éléments de notre matrice matrice .

4.4.1.2 La moyenne arithmétique

Elle retourne la moyenne arithmétique de l'échantillon des données qui peuvent être une séquence ou un itérateur.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

La moyenne arithmétique est la somme des données divisée par le nombre de points de données. On l'appelle communément "la moyenne", bien qu'il ne s'agisse que d'une des nombreuses moyennes mathématiques différentes. C'est une mesure de l'emplacement central des données.

4.4.1.3 La variance

En statistique, la variance décrit comment les valeurs d'un échantillon sont distribuées,

Elle représente la différence entre la moyenne des carrés et le carré de la moyenne.

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

4.4.1.4 L'écart-type

Il est défini comme la racine carrée de la variance .Il se note en général avec la lettre grecque σ .

$$\sigma = \sqrt{V}$$

4.4.1.5 La médiane

En probabilités, la médiane est la valeur qui nous permet de diviser une population en deux parties égales.

Si les observations sont rangées en ordre croissant, une partie inférieure et une partie supérieure, elle représente le milieu des valeurs.

$$M = \frac{\frac{x_n}{2} + \frac{x_{n+1}}{2}}{2}$$

4.4.1.6 La médiane haute

Retourne la médiane élevée des données. Si les données sont vides, `StatisticsError` est affiché. Les données peuvent être une séquence ou un itérateur.

La médiane élevée est toujours un membre de l'ensemble de données. Lorsque le nombre de points de données est impair, la valeur centrale est retournée. Lorsqu'elle est paire, la plus grande des deux valeurs centrales est retournée.

4.4.1.7 La médiane basse

Retourne la médiane inférieure des données numériques. Si les données sont vides, `StatisticsError` est affiché. Les données peuvent être une séquence ou un itérateur.

La médiane inférieure est toujours un membre de l'ensemble de données. Lorsque le nombre de points de données est impair, la valeur centrale est retournée. Lorsqu'elle est paire, la plus petite des deux valeurs centrales est retournée.

4.4.1.8 Le taux de passage a zéro

'Le Zero Crossing Rate (ZCR) en anglais'

Est le taux de changement de signe d'un signal.

4.5. RÉSEAUX NEURONES POUR LA CLASSIFICATION 21

On l'utilise pour connaître la transition du signal de la partie positive vers la négative et vice-versa. Il est défini par : [9]

$$Z_{SQI} = - \sum_{n=1}^N I\{y < 0\}$$

On a calculé, combien de fois le signal passe par le 0, on a trouvé que généralement les bons signaux passent par le 0 au moins 1 fois ou plusieurs, par contre les mauvais ne passent pas par la ligne des abscisses.

4.4.1.9 L'asymétrie (skewness)

Calcule l'asymétrie d'un ensemble de données.

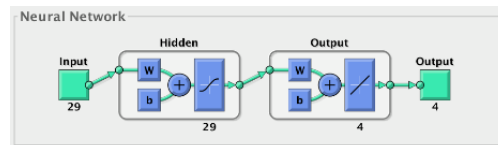
Une valeur d'asymétrie > 0 signifie qu'il y a plus de poids dans la queue gauche de la distribution.

La fonction skewtest peut être utilisée pour déterminer si la valeur d'asymétrie est suffisamment proche de 0, statistiquement parla

$$1/N \sum_{i=1}^N [x_i - \frac{\sum x}{N}]^3$$

4.5 Réseaux neurones pour la classification

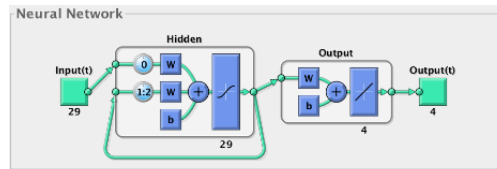
Feed Forward Le réseau neuronal à effet de rétroaction a été le premier et le plus simple type de neurones artificiels. réseau conçu. Dans ce réseau, l'information se déplace dans une seule direction, vers l'avant, depuis les nœuds d'entrée jusqu'aux nœuds de sortie, en passant par les nœuds cachés. Il n'y a pas de cycles ou de boucles dans le réseau. Cela signifie que chaque couche suivante a la connexion à la couche précédente. La dernière couche produit la sortie de la fonction réseau neuronal.



LSTM Les réseaux de mémoire à long terme (LSTM) sont bien adaptés à la classification, au traitement et à l'analyse des données. des séries chronologiques, puisqu'il peut y avoir des décalages dans le temps. de durée inconnue entre les événements importants d'une série chronologique. Les LSTMs étaient de l'explosion et de la disparition des problèmes de gradient

qui peuvent causer des problèmes de être rencontrés lors de la formation des réseaux traditionnels récurrents. Le programme mis en œuvre LSTM a une couche d'entrée de séquence avec 29 caractéristiques suivie d'une couche LSTM avec 290 unités cachées, dans lesquelles le mode de sortie est une séquence. Après celle-ci, nous avons une autre couche LSTM avec 174 unités cachées. Il a une couche de connexion complète avec 4 en raison du nombre de classes, suivi d'un calque max soft et d'un calque de type une couche de classification.

Recurrent Similaire à Feed Forward à l'exception près que chaque couche a une valeur récurrente avec un retard associé. Cela permet une réponse dynamique infinie à les données d'entrée des séries chronologiques.



CNN Convolutional neural network (CNNs) utilisent relativement peu de prétraitement par rapport à d'autres algorithmes de classification d'images. Cela signifie que le réseau apprend le qui, dans les algorithmes traditionnels, étaient conçus à la main. Cette indépendance des connaissances préalables et de l'effort humain dans la conception des caractéristiques est un avantage majeur. Le Cnn implémenté a une couche convolutionnelle avec une taille de piscine de 5x5, ayant 29 filtres, avec un rembourrage de 2, suivis d'une couche d'accumulation maximale avec une taille de regroupement 3x3 et foulée de 2 et un relulayer. La deuxième couche convolutionnelle a la même la taille de la piscine, le rembourrage et les filtres de la précédente, et son suivi d'une re-doublure. et une couche de piscine moyenne d'une taille de piscine de 3x3 et d'une foulée de 2. La couche convolutionnelle a également une taille de regroupement 5x5, 2 rembourrages, mais a 58 filtres, suivi par une couche de re-doublure et une couche de pooling moyenne avec une taille de piscine de 2x2 et une foulée de 2. Ensuite, nous avons une couche de connexion complète avec 58 sorties, suivie d'une couche de re-doublure. et une autre couche entièrement connectée avec 4 sorties, ceci parce que le nombre de sorties doit être égal au nombre de classes de l'ensemble de données, suivi d'une couche de matelas souple. et une couche de classification.

K-Nearest Neighbors

Définition : En intelligence artificielle, le K-NN est une méthode utilisé à des fins de classification ou de régression. Son fonctionnement est

inspiré de la notion d'amitié et de voisinage dans la vie réel : “dis-moi qui sont tes amis, je te dirais qui tu es...”.

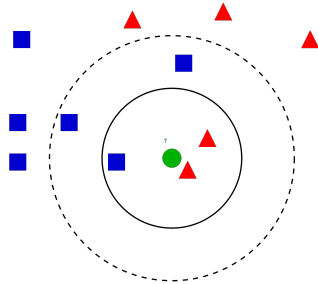


FIGURE 4.1 – K-Nearest Neighbors

Fonctionnement de KNN : Le fonctionnement de K-NN a besoin de 3 choses principaux : Un ensemble de données \mathbf{D} , une fonction de définition distance \mathbf{d} , un nombre entier \mathbf{K} . Il calcule toutes les distances entre la nouvelle donnée et les autres données, après il prend le nombre de \mathbf{K} comme référence, par exemple si le nombre $\mathbf{K} = 3$, il prend juste les 3 données les plus proches de la nouvelle donnée, ensuite il effectue la nouvelle donnée à la classe qui a plus d'observation.

Comment choisir la valeur \mathbf{K} : Le choix de la valeur \mathbf{K} joue un rôle très important dans la classification par K-NN. Si le nombre de voisins est petit, on va avoir un sous apprentissage .En outre, si on utilise un nombre de voisins grand, on risque d'avoir un sur-apprentissage et donc une mauvaise généralisation sur des observations non déjà vu. Ainsi, il faut trouver la meilleure valeur de \mathbf{K} pour que notre classification sera efficace.

On a proposé comme solution de dessiner un graphe qui montre la variation de l'erreur moyenne en fonction de la valeur de \mathbf{k} , la meilleure valeur de \mathbf{K} sera l'intervalle où on voit une diminution de l'erreur moyenne.

4.6 Resultat et Conclusion

Pour cette expérimentation, nous avons utilisé Matlab, sur lequel nous avons implémenté plusieurs scripts pour modifier la base de données sélectionnée. L'équilibrage de classe s'assurera que l'élément nombre de points de différents états sera le même de points Ictal, par exemple, si nous avons 100 points Ictal, alors l'Inter-Ictal, le Pré-Ictal et le Pos-Ictal ont 100 points Ictal. 100 points chacun. Pour voir les performances des réseaux de neurones, ont

été faites plusieurs tests avec la formation 0les tests de 0et ont été analysés les attributs suivants :

- - SE, sensibilité
- - SP, spécification
- - PIA, précision pré-ictale
- - IA, Précision de l'Ictal
- - AC, Précision générale

Resultats CNN

SE	SP	PIA	IA	AC%
77%	57%	50%	76%	25
%				
75%	60%	52%	76%	32
%				
58%	61%	33%	76%	33
%				
82%	58%	39%	76%	26
%				
68%	60%	42%	88%	32
%				

Resultats LSTM

SE	SP	PIA	IA	AC%
67%	58%	34%	82%	25%
69%	57%	37%	96%	24%
62%	58%	31%	96%	22%
74%	62%	36%	91%	35%
64%	61%	43%	85%	32%

Le meilleur réseau neuronal a été réalisé avec 54802 it, et la raison possible est la suivante que c'est l'ensemble de données avec le plus de phases crises. L'autre patient avait une très mauvaise en raison du faible nombre de crises. A été vérifié que l'ensemble de données le traitement est très important, et aussi, pour avoir un bon classificateur nous devons avoir un long ensemble de données, pour fournir les meilleurs résultats possibles. C'est

vraiment dur d'avoir de bons pour la prédiction et la détection des crises épileptiques. Comme on peut le voir dans les résultats fournis ,les meilleurs réseaux neuronaux sont CNN et LSTM. En termes de les valeurs les plus élevées apparaissent sur CNN (tableau 2), avec une valeur maximale de 82 %, et une moyenne de 5 ce réseau neuronal présente également les meilleures valeurs en termes de prédiction (précision pré-crise). Dans la détection des crises Le LSTM est celui qui a les valeurs les plus élevées, atteignant une efficacité maximale de 6

Enfin, la meilleure façon de prédire et de détecter une crise est de s'assurer que le Le réseau neuronal utilisé a été formé avec le même type de crise que le patient dans les cas suivants question.

Chapitre 5

Quelques extraits de code

```
    signal avec crise
clc;
clear all;
close all;
selection d'electrode utilisÈ pour la detetction de peaks pour la periodes de
crise
    chb0802 = load('chb0802_data.txt');
    chil82chn21 = chb0802(:, 22); FT9_FT10
    chb0805 = load('chb0805_data.txt');
    chil85chn21 = chb0805(:, 22);
    chb0811 = load('chb0811_data.txt');
    chil81chn21 = chb0811(:, 22);
    chb0813 = load('chb0813_data.txt');
    chil813chn21 = chb0813(:, 22);
    chb0821 = load('chb0821_data.txt');

    chil821chn21 = chb0821(:, 22); FT9_FT10

    Start_time2 = 2670 * 256;
    End_time2 = 2841 * 256;
    [crsch82chn21, crssglch82chn21] = extra_crise(chil82chn21, Start_time2, End_time2);

    Start_time5 = 2856 * 256;

    End_time5 = 3046 * 256;

    [crsch85chn21, crssglch85chn21] = extra_crise(chil85chn21, Start_time5, End_time5);
```

```

    StartTime11 = 2988 * 256;
    EndTime11 = 3122 * 256;
    [crsch81chn21, crssglch81chn21] = extracrise(chil81chn21, StartTime11, EndTime11);

    StartTime13 = 2417 * 256;
    EndTime13 = 2577 * 256;
    [crsch813chn21, crssglch813chn21] = extracrise(chil813chn21, StartTime13, EndTime13);

    StartTime21 = 2083 * 256;
    EndTime21 = 2347 * 256;
    [crsch821chn21, crssglch821chn21] = extracrise(chil821chn21, StartTime21, EndTime21);
    reshapeperiodedecriseen8seconds
    [recrsch82chn21] = reshapecrs(crsch82chn21);
    [recrsch85chn21] = reshapecrs(crsch85chn21);
    [recrsch811chn21] = reshapecrs(crsch811chn21);
    [recrsch813chn21] = reshapecrs(crsch813chn21);
    [recrsch821chn21] = reshapecrs(crsch821chn21);

```

Chapitre 6

Références

Andrea Němcová, MIT-BIH Arrhythmia Database, Jan. 9, 2018,<https://www.physionet.org/content/pwave/1.0.0/>.

Shayan Fazeli, ECG Heartbeat Categorization Dataset, 2018-05-31,<https://kaggle.com/shayanfazeli/heartbeat>

Zemzemi, N. (2009). Étude théorique et numérique de l'activité électrique du cœur : Applications aux électrocardiogrammes (Doctoral dissertation, Paris 11).

Hussain, H., Fatt, L. L. (2007, December). Efficient ECG signal classification using sparsely connected radial basis function neural network, Control and Signal Processing (pp. 412-416).

Beaufils, B. (2000). Statistiques appliquées à la psychologie : statistiques descriptives (Vol. 1). Editions Bréal.

Slimani, Y., Essegi, M. A., Samb, M. L., Camara, F., Ndiaye, S. (2014). Approche de sélection d'attributs pour la classification basée sur l'algorithme RFE-SVM.

Philippe Beraud - MSFT, Evaluer un modèle en apprentissage automatique, June 17, 2014,https://blogs.msdn.microsoft.com/big_data_france/2014/06/17/evaluer-un-modle-en-apprentissage-automatique/.