

# Prosod'IA : AIDE AU DIAGNOSTIC PRECOCE DES TSA

## RAPPORT DE PROJET

ANAEILLE MARIETTE, LIVIO DANINTHE, TOM LALIEU,  
NATHAN LECOIN, LOUKA MORANDI, LILA RAZANI

RESPONSABLE : NADIA MADAOUI

DATE DU PROJET : 25 JUIN 2024



## Remerciements

Nous tenons à remercier Mme Madaoui, pour nous avoir proposé ce projet très enrichissant, ainsi que pour nous avoir guidés et conseillés tout au long de la mise en œuvre du projet. Nous remercions également Sandrine Sonié, qui a accepté de nous donner de son temps, sans quoi nous n'aurions pas pu souligner des questions essentielles. Grâce à cet échange, nous pourrions par ailleurs envisager un partenariat futur, à condition que notre projet soit fiable. Enfin, nous tenons à adresser nos remerciements à ESIEE Paris qui nous a donné l'opportunité de mener à bien un projet technique, qui nous a permis de nous rendre compte des futures responsabilités d'un ingénieur.

<b>Remerciements</b>	<b>1</b>
<b>Introduction</b>	<b>3</b>
<b>Chapitre I : Présentation du projet</b>	<b>4</b>
Partie 1 : Cadre du projet	4
Partie 2 : Contexte du projet	4
2.1- Origine des TSA et conditions caractéristiques	4
2.2- Processus actuel de diagnostic	5
Partie 3 : Méthodologie	6
<b>Chapitre II : Spécification des besoins</b>	<b>8</b>
Partie 1 : Répartition des tâches	8
Partie 2 : Ressources utilisées	8
<b>Chapitre III : Réalisation</b>	<b>9</b>
Partie 1 : Extraction des caractéristiques	9
1.1- Documentation	9
a- La prosodie	9
b- MFCC, spectrogrammes, scalogrammes, mel-spectrogramme, pitch	11
1.2- Filtrage de la voix	12
1.3- Génération des caractéristiques sous forme d'image	13
1.4- Etude comparative	15
Partie 2 : L'IA pour l'aide au diagnostic	17
2.1- Choix de l'architecture de l'IA	17
2.2- Etapes préliminaires	19
a- Présélection des réseaux de neurones pré-entraînés	19
b- Prétraitement des données	20
2.3- Tests et entraînements	21
a- Réalisation des tests sur Matlab	21
b- Réalisation des tests sur Python	22
Partie 3 : Développement de l'application d'aide au diagnostic	24
<b>Conclusion technique</b>	<b>29</b>
<b>Bilan du projet</b>	<b>30</b>
Partie 1 : Référentiel compétences de l'ESIEE	30
1.1- Concevoir des systèmes, applications ou solutions d'ingénierie complexes	30
1.2- Réaliser des systèmes, applications ou solutions d'ingénierie complexes	30
1.3- Gérer un projet	31
1.4- Agir avec une démarche scientifique et éthique	32
Partie 2 : Bilan personnel	32
2.1- Retour d'expérience de Lila	32
2.2- Retour d'expérience d'Anaelle	33
2.3- Retour d'expérience de Nathan	33
2.4- Retour d'expérience de Louka	34
2.5- Retour d'expérience de Livio	34
2.6- Retour d'expérience de Tom	35
<b>Bibliographie</b>	<b>36</b>
<b>Annexe</b>	<b>37</b>

## Introduction

Aujourd’hui en France, un enfant sur cent est atteint de Trouble du Spectre Autistique (TSA), un trouble neurodéveloppemental qui touche principalement la communication socio-émotionnelle. La détection des TSA est encore un processus long et complexe, faisant intervenir différents professionnels de santé et de multiples questionnaires. Le diagnostic peut ainsi s’étendre sur plusieurs mois, voire plusieurs années.

Cependant, les professionnels manquent d’outils d’évaluation et les études sont difficiles à mener à cause de la grande variabilité interindividuelle des symptômes. Or, diagnostiquer dès le plus jeune âge afin de prendre en charge le plus tôt possible les patients est fondamental dans le but de suivre au mieux l’enfant et de lui proposer les outils les plus adaptés.

Actuellement, il y a 700 000 cas recensés en France, dont 100 000 de moins de 20 ans. Ces personnes vivent un handicap social, allant d’anomalies comportementales (telles que la réticence au changement et une tendance à la répétition de comportements ou de discours), jusqu’à des altérations dans la capacité à établir des interactions sociales et à communiquer.

Ainsi, pour tenter de comprendre les troubles dans le développement, et améliorer les diagnostics actuels, de plus en plus de chercheurs se réfèrent aux procédés d’acquisition du langage. C’est à partir de ces recherches que notre projet réside dans la conception d’une aide au diagnostic précoce des TSA, basée sur l’étude des caractéristiques vocales, avant l’acquisition du langage, notamment sur l’étude de la prosodie.

Dans un premier temps, nous préciserons certaines notions importantes sur les TSA, et définirons la prosodie et ses caractéristiques à exploiter. Dans un second temps, nous présenterons les outils techniques utilisés pour analyser la base de données, et les stratégies auxquelles nous avons dû faire face pour mener à bien ce projet.

## Chapitre I : Présentation du projet

Dans ce chapitre, il est question d'introduire le sujet du projet. Nous définirons les TSA dans leur globalité et le contexte du projet, c'est-à-dire les recherches actuelles et ce que nous pouvons y apporter. Nous présenterons également la méthodologie que nous avons mise en place pour mener à bien notre projet.

### Partie 1 : Cadre du projet

Dans le cadre de notre première année de cycle ingénieur, nous avons été chargés de réaliser un projet d'une durée de sept semaines. Initialement, nous étions vus être attribués le projet *Analyse de signaux ECoG pour la cartographie spatio-temporelle des aires du langage*, en collaboration avec l'INSERM. L'objectif était de mettre en place une cartographie des zones du cerveau qui s'activent lors de la parole. Des analyses en amont sur des cochons ont été effectuées pour enregistrer leurs vocalises, cependant l'INSERM n'a pas pu nous fournir cette base de données à temps. C'est pourquoi nous n'avons pas pu continuer de travailler sur ce projet, bien que nous fussions très intéressés par celui-ci.

Après réflexion avec notre tutrice de projet, Mme Madaoui, nous avons décidé de changer de sujet. Si nous avions poursuivi, nous aurions continué le projet initial avec des ECG. Or, les résultats auraient été bien moins pertinents et intéressants. Nous nous sommes donc rabattus sur notre projet actuel, qui a immédiatement attiré notre attention.

Notre équipe de projet se compose de Livio Daninthe, Tom Lalieu, Nathan Lecoin, Anaelle Mariette, Louka Morandi et Lila Razani.

### Partie 2 : Contexte du projet

#### 2.1- Origine des TSA et conditions caractéristiques

Les TSA sont d'origine multifactorielle et largement génétique. La diversité des troubles et les capacités d'insertion sociale sont très variables selon les personnes. Par ailleurs, c'est pour cette raison que dans la suite du rapport, nous n'employons pas le terme d'autisme, mais de patients atteints de TSA.

En effet, il existe différents syndromes tels que le syndrome de Rett, qui correspond à une déficience intellectuelle, le syndrome d'Asperger, correspondant à un très bon développement intellectuel ou encore le syndrome de l'X fragile. La variabilité de ces syndromes est ainsi à l'origine de la difficulté actuelle à établir un diagnostic.

Chacun de ces syndromes résulte d'anomalies du neurodéveloppement. En particulier, elles correspondent à des défauts de mise en place et d'organisation de certains réseaux cérébraux spécialisés (dédiés à la communication sociale et à la modulation du comportement en fonction de l'environnement). À l'échelle de la biologie cellulaire, cela se traduit par une mutation de plusieurs

centaines de gènes responsables de formation du système nerveux et des connexions synaptiques, ainsi qu'à la synthèse de substances chimiques indispensables au bon fonctionnement du cerveau.

Ces mutations peuvent donc être d'origine génétique, ou bien avoir lieu pendant la grossesse, suite à une exposition à des facteurs de risque, tels que des médicaments comme le valproate ou des toxines comme l'alcool.

Les conditions caractéristiques d'un individu atteint de TSA résident dans l'altération des interactions sociales, avec des problèmes de communication de langage et non verbale. Ces conditions comportent aussi des troubles du comportement comme des intérêts et activités restreintes, stéréotypés et répétitifs, et des réactions sensorielles inhabituelles.

Les troubles affectant les relations et la communication sont des caractéristiques fondamentales des TSA. En pratique, ils se définissent par la difficulté à établir les contacts nécessaires à la construction d'une relation interpersonnelle. Adultes et enfants ont des difficultés à comprendre les émotions des autres, et à utiliser les gestes, les expressions du visage, le regard ou le ton de la voix.

Ces troubles se définissent également par des comportements répétitifs (balancements du corps, battements des mains, tournoiements ...), auto-agressifs (se mordre les mains, se cogner la tête...) ou inappropriés (pleurer ou rire sans raison apparente...). De plus, ils sont à l'origine d'une faible tolérance au changement, où une situation imprévisible peut provoquer une réaction d'angoisse ou de panique, de colère ou d'agressivité.

Enfin, les TSA s'accompagnent souvent de troubles du sommeil et de l'apprentissage, de dépression, d'anxiété, d'hyperactivité et un individu sur cinq souffre également d'épilepsie.

## 2.2- Processus actuel de diagnostic

En principe, le processus de diagnostic des TSA débute dès les premières visites médicales. Lors de chaque visite de 0 à 6 ans, le professionnel de santé chargé du suivi habituel de l'enfant examine la communication, la motricité, le langage et cherche à identifier des signes d'alertes de TSA. Ce processus est par la suite renforcé s'il y a une apparition de symptômes ou une préoccupation de la part des parents concernant le développement de leur enfant. En particulier en ce qui concerne la communication sociale et le langage ou lorsqu'il y a une régression des habiletés langagières ou relationnelles.

Les signes majeurs de TSA à observer chez les jeunes enfants dépendent de l'âge de ce dernier. À partir de 12 mois, il faut observer l'absence de babillage ou de gestes de communication. Dès 18 mois, les parents doivent rester vigilants à l'égard de l'absence de parole et à partir de 24 mois à l'absence de formulation de phrase.

D'autres signes d'alerte de TSA, à partir de 18 mois, incluent la persistance de particularités de développement de l'enfant telles que son niveau de vigilance, la régulation des émotions ou la diversification alimentaire.

Lorsque ces signes sont remarqués par les parents ou les enseignants, il est nécessaire d'orienter l'enfant vers le médecin traitant. Ainsi, le diagnostic débute suite à l'inquiétude des parents ou des enseignants. Il peut également débuter si l'enfant est exposé à un risque plus élevé de TSA que la norme, en particulier s'il y a une hérédité génétique.

Au cours de la visite médicale de contrôle, a lieu un examen clinique approfondi sur le développement de l'enfant. Le médecin s'appuie sur des outils de repérage en fonction de l'âge de l'enfant comme M-CHAT, pour ceux de 16 à 30 mois.

Toutes les étapes de ce diagnostic sont représentées par un schéma récapitulatif. [cf. annexe 1]

De nouveaux outils sont actuellement en cours d'étude afin de détecter les TSA depuis chez soi, dans le cadre habituel de vie de l'enfant. Cela permettrait à l'enfant de ne pas faire ce dépistage dans un environnement étranger, comme les cabinets médicaux ou les hôpitaux.

Ces instruments incluent des analyses automatiques de films ou même des suivis du regard. La linguistique est aussi mise à contribution pour approfondir l'analyse des caractéristiques formelles du langage dans les TSA. Ces nouvelles recherches présentent des similitudes avec notre outil, qui pourrait également être employé en dehors d'un contexte médical.

À ce jour, notre outil ne sera ouvert qu'à des professionnels de santé pour des raisons éthiques, afin d'éviter certaines dérives. Cependant, si notre projet est suffisamment fiable et fonctionnel, nous pourrions considérer la perspective de le rendre accessible à des particuliers.

Malgré ce processus de diagnostic qui pourrait être considéré comme suffisant, en réalité, le diagnostic des TSA est encore effectué beaucoup trop tardivement, au plus tôt vers l'âge de trois ou quatre ans, mais souvent bien plus tard. Or, la prise en charge précoce de l'enfant est une garantie pour le bon développement cognitif de l'enfant.

En somme, nous pouvons constater que les diagnostics disponibles à ce jour sont limités en outils de repérage, peuvent prendre des mois à être confirmés et sont effectués beaucoup trop tard. Nous avons donc choisi ce sujet en raison de l'urgence dans laquelle se retrouvent les personnes concernées, face au manque de mesures prises à jour.

Dans ce contexte, on peut donc affirmer la pertinence et l'efficacité de notre outil d'aide au diagnostic précoce des TSA. En effet, notre outil offre aux médecins de première ligne une possibilité supplémentaire pour détecter les TSA. Les professionnels seront non seulement épaulés par des questionnaires, mais également par une analyse scientifique approfondie des vocalises des jeunes enfants. Toutefois, notre outil reste complémentaire aux pratiques actuelles. Ainsi, il ne faut pas l'isoler mais prendre également en compte les outils de diagnostics actuels.

### Partie 3 : Méthodologie

Afin de réaliser le projet rigoureusement, nous avons dû suivre une certaine méthodologie et des étapes bien définies.

Avant toutes choses, le sujet proposé ne faisant pas partie d'un domaine que nous maîtrisons, nous avons donc pris le temps de nous documenter. Ainsi, la première étape était principalement régie par la bibliographie, afin de déceler les moyens existants et accessibles pour mener à bien le projet. Cette phase de bibliographie nous a également permis de mieux comprendre les logiciels et les technologies que nous utilisons par la suite.

Cela conduit ainsi à la seconde partie, définie par l'analyse d'enregistrements de vocalises de bébés de 0 à 24 mois à partir d'une base de données. Cette dernière a été récupérée par notre responsable, Mme

Madaoui, sur des enregistrements audio ou des vidéos en libre-service en ligne. Cependant, au fil des semaines, nous nous sommes rendus compte que cette base de données n'était pas optimale.

Pour des raisons d'autorisation, nous n'avions pas accès à une base de données normée. Afin d'être éclairés sur la question, nous avons contacté Sandrine Sonié, coordinatrice du Centre de ressources autisme de Rhône Alpes.

Lors de notre échange, nous avons été mis en garde sur le protocole de diagnostic. Il doit passer par un comité de protection des personnes, ce qui nous permettrait d'être sûrs que les familles aient bien donné leur accord pour utiliser les enregistrements à des fins de recherche.

De surcroît, Sandrine Sonié estime que pour assurer la fiabilité et la validité du diagnostic, nous devons nous renseigner sur la fiabilité et la validité de notre base de données. En effet, si nous ne prenons pas connaissance des conditions dans lesquelles ont été diagnostiqués les enfants de notre base de données, en s'assurant que le protocole est fiable et encadré, alors on ne pourrait pas prétendre à une application réellement fonctionnelle.

Toutefois, n'ayant pas les dispositions nécessaires pour se baser sur des enregistrements encadrés, nous avons décidé de garder cette base de données, et de continuer à entraîner notre IA avec celle-ci.

En dépit de cette problématique, pour pouvoir travailler sur la base de données, nous avons décidé de regrouper les enregistrements en deux catégories : sains et atteints de TSA. Puis, nous les avons répartis en sous-catégories, d'abord par sexe, puis par âge, et enfin par activité (un exemple serait une fille saine, entre 6 et 12 mois, en train de pleurer).

Après avoir trié la base de données, il a fallu l'analyser. Pour cela, nous avons dans un premier temps généré plusieurs images telles que les MFCC ou les spectrogrammes, que l'on définira dans la suite du rapport, à l'aide de fonctions existantes sur Matlab. Dans un second temps, nous avons récupéré ces caractéristiques et établi un tableau comparatif afin de déterminer quelles images seraient exploitables.

Enfin, la troisième et dernière partie du projet réside dans l'entraînement de l'IA, à partir de réseaux de neurones pré-entraînés. Ce travail de Transfert Learning prend en paramètres les images générées précédemment, et en décelant les caractéristiques de chaque cas, il est capable de classifier les patients sains et ceux atteints de TSA.

Afin d'être le plus fiable possible, la complexité réside en particulier dans le choix du réseau de neurones le plus précis et adapté. De plus, pour entraîner au mieux l'IA, nous avons réparti la base de données en 80% pour les entraînements et 20% pour les tests.

Une fois les tests validés, nous convertirons le code en un exécutable, plus facile d'accès pour les utilisateurs.

En résumé, nous avons établi un organigramme afin de regrouper l'ensemble des tâches. [cf. annexe 4]

Pour conclure ce premier chapitre, après avoir pris connaissance des recherches actuelles sur les TSA et des outils disponibles pour les diagnostiquer, nous pouvons affirmer que notre application est pertinente et qu'elle représente une véritable avancée dans le diagnostic précoce des TSA. Si nous parvenons à la rendre fiable et opérationnelle, elle permettrait donc d'offrir de nouvelles opportunités aux patients.

Pour y parvenir, nous passerons par deux grands chapitres techniques, le premier étant la génération des caractéristiques vocales en tant qu'images, le second étant l'utilisation de ces images grâce au Transfert Learning.

## Chapitre II : Spécification des besoins

Ce chapitre évoque les besoins auxquels nous avons eu recours lors de ce projet. Il détaille ainsi la répartition des tâches au sein des membres du groupe et les ressources utilisées, à la fois matérielles et humaines.

### Partie 1 : Répartition des tâches

Les tâches du projet ont été divisées en trois catégories distinctes. Afin de mener à bien le projet et de le terminer à temps, chaque membre s'est vu attribuer une catégorie en binôme.

Ainsi, le projet repose premièrement sur la compréhension des TSA et des recherches sur les propriétés des images telles que les MFCC et les spectrogrammes. Deuxièmement, il se base sur la réalisation de ces caractéristiques. Troisièmement, il consiste sur la mise en place d'une IA qui analyse ces images pour enfin détecter la présence de TSA.

Les trois sujets ont donc été divisés de la sorte : Anaelle Mariette et Lila Razani ont été chargées de la compréhension des TSA et de recherches sur les images. Ensuite, Livio Daninthe et Tom Lalieu étaient responsables de la partie consacrée à la réalisation des images. Enfin, l'IA a été réalisée par Louka Morandi et Nathan Lecoin.

Malgré ces répartitions, quelques tâches ont été réalisées collectivement car certains sujets se sont chevauchés. En outre, le code Matlab fournissant toutes les caractéristiques nécessaires (MFCC, spectrogramme, mel-spectrogramme ou encore scalogrammes) pour chaque piste audio de notre base de données a été réalisé par Anaëlle, Tom, Livio et Lila. Ceci a donc permis de fournir ces images à Nathan et Louka pour les entrer en paramètre dans l'IA, et ainsi les utiliser pour son apprentissage.

### Partie 2 : Ressources utilisées

Lors de la mise en œuvre du projet, nous avons dû mobiliser diverses ressources, à la fois des ressources humaines et matérielles.

Pour ce qui est de la gestion du projet, nous avons tout d'abord utilisé un GANTT afin de garder des objectifs précis et une vision globale du projet. De plus, nous avons suivi les tâches de chacun en temps et en heure à l'aide d'un Trello, ce qui nous a permis de suivre l'avancée du projet.

En ce qui concerne les ressources matérielles, nous avons dû recourir à plusieurs logiciels, tels que TensorFlow, Matlab et Visual Studio Code.

La première approche a été d'utiliser Matlab afin de développer les premiers codes pour récupérer les caractéristiques vocales.

Toutefois, nous avons dû changer de logiciel dans le but de faciliter la création de l'application et d'harmoniser l'intégralité du projet. Nous avons ainsi opté pour l'utilisation de Python, autant pour les codes initialement développés sur Matlab que pour la programmation de la partie IA.

Pour ce faire, nous nous sommes dirigés vers l'outil open source TensorFlow.

Concernant les ressources humaines, nous avons sollicité plusieurs personnes. Tout d'abord, notre tutrice de projet, Mme Madaoui, nous a aiguillé et épaulé tout au long de la réalisation du projet. En effet, elle nous a fourni toutes les ressources nécessaires au début du projet, telles que la bibliographie et la base de données. Elle était également disponible lorsque nous avions besoin de plus de précisions concernant des sujets complexes, par exemple pour des explications supplémentaires des caractéristiques obtenues ou sur les réseaux de neurones à exploiter. Nous avons aussi fait appel à Sandrine Sonié pour avoir un avis extérieur sur notre projet [cf. annexe 3].

Pour conclure le second chapitre, ce projet a été réparti naturellement selon les compétences de chacun, en prenant compte de ce que les membres du groupe peuvent apporter individuellement. Par ailleurs, c'est un projet qui n'a pas nécessité de ressources matérielles particulières, étant constitué d'un travail d'analyse de données et de programmation.

## Chapitre III : Réalisation

Ce chapitre regroupe toute la partie technique qui a permis de développer le projet. Il s'articule selon les trois thèmes de conception de l'aide au diagnostic.

En premier lieu, nous abordons comment nous avons extrait les caractéristiques vocales de la voix d'un bébé. D'abord en cherchant les informations nécessaires et en s'informant, puis en les extrayant sous forme d'images.

En second lieu, grâce à l'étape précédente, nous récupérons donc ces images afin de les tester avec l'IA. C'est en choisissant le bon réseau de neurones et les bons paramètres que nous sommes parvenus à différencier les images des enfants sains de celles des enfants atteints de TSA. Une fois que notre IA était opérationnelle, nous avons pu passer à la dernière étape.

En troisième lieu, nous avons ainsi développé une application dédiée aux professionnels de santé, afin qu'ils puissent utiliser notre outil à des fins cliniques. Cette application est donc totalement automatisée et facile d'utilisation, pour les aider au mieux à diagnostiquer les enfants concernés.

### Partie 1 : Extraction des caractéristiques

#### 1.1- Documentation

##### a- La prosodie

Basant notre projet sur l'analyse des caractéristiques vocales, la prosodie constitue une approche intéressante à étudier. Par définition, elle décrit les variations de hauteur, d'intensité et de durée qui correspondent aux phénomènes d'intonation, d'accentuation, de débit ou encore la perception des pauses. C'est le premier système fonctionnel précoce et donc l'une des principales voies d'accès au langage pour les enfants.

Plusieurs travaux ont mis en évidence le rôle fondamental de la prosodie pendant la période préverbale. Selon l'étude sur la prosodie en acquisition de Karine MARTEL, observer l'évolution de la prosodie permettrait de déceler à quel moment le système langagier devient incohérent ou présente des aspects atypiques.

Actuellement, les professionnels manquent d'outils d'évaluation et comme mentionné précédemment, les études sont difficiles à mener à cause de la grande variabilité interindividuelle des symptômes.

Ainsi, l'objectif du projet est de fournir un nouvel outil, se concentrant sur d'autres indices communicationnels importants tels que la gestion des pauses, les tours de parole et l'alliance d'indices linguistiques.

Ces indices ne sont à ce jour pas l'objet d'études concrètes, or les formes précoce du langage portent en elles les germes des troubles : c'est à ces formes précoce qu'il faut s'intéresser, c'est-à-dire aux gazouillis, qui se transforment ensuite en vocalisations. Il faut donc pour cela réunir différentes approches autour du traitement du signal.

Par ailleurs, c'est à partir de 3 mois que les émissions vocales prennent un caractère différencié. Aux alentours de 4 mois, le bébé peut ensuite mieux moduler et contrôler les changements de hauteur, de durée et d'intensité de sa voix.

L'absence ou l'exagération de l'allongement de la syllabe finale peut donc être considérée comme un marqueur potentiel d'un développement standard ou retardé.

Tous les travaux effectués jusqu'ici montrent que la prosodie constitue bel et bien une base pour l'avènement de la grammaire.

Il reste ainsi plusieurs indices à explorer dans la parole du très jeune enfant, comme le volume, le débit et les changements de tonalité.

Lorsque la prosodie est prise en compte dans des tests ou échelles d'évaluation, les principales caractéristiques observées sont réduites aux possibilités d'imitation de séquences langagières ou au fait d'utiliser une gamme de fréquence aux différentes amplitudes qui va du murmure aux cris suraigus.

Les vocalisations sont très peu observées. Il n'est fait aucune mention de leur nature, de leur forme, de leurs régulations fréquentielles. Or la prise en compte de ces indices permettrait certainement d'enrichir l'interprétation des résultats obtenus.

L'analyse des aspects prosodiques pourrait de la même façon compléter les bilans obtenus.

En effet, il est nécessaire de préciser que les analyses que l'on développe au cours du projet ne se suffisent pas à elles-mêmes et c'est pourquoi nous prenons les précautions de fournir uniquement une aide au diagnostic, avec l'intervention systématique d'un professionnel de santé. L'application est donc un indice sur l'état de santé de l'enfant, mais n'a pas pour objectif de remplacer le rôle du professionnel. Notre facteur de risque sorti d'un contexte développemental et d'éléments cliniques n'a pas de valeur.

En 1997, Gabrielle Konopczynski et Shirley Vinter mettent en effet en avant que les enfants sourds, une fois appareillés et en mesure de contrôler la qualité de leur voix (la cohérence de la structure mélodique), améliorent leur apprentissage et leur compréhension des règles pragmatiques de la communication sociale.

En l'occurrence, les enfants atteints d'autisme ont des difficultés significatives avec le module prosodique, comme une impossibilité ou une grande gêne à exprimer et comprendre des états affectifs, ou encore à utiliser l'accent d'insistance ou à reproduire des phrases courtes ou longues.

Les troubles de la prosodie expressive engendrent une barrière communicative pour ces enfants. Ce qui provoque une altération de la boucle perception/production.

Il faut donc isoler le rôle de chaque paramètre grâce à l'utilisation de logiciels d'analyse du signal performant, et recueillir des données brutes au domicile des sujets ou au sein d'institutions spécialisées de façon régulière (de qualité de son optimale, dans des situations de communication variées, incluant des partenaires variés). Chaque sujet peut ainsi donner lieu à la description d'un profil. Cette démarche est obligatoire pour que l'étude de la prosodie et de son développement puissent contribuer à une approche clinique.

En somme, on peut poser l'hypothèse que les « anomalies » prosodiques sont à considérer comme de potentiels marqueurs précoce de trajectoires déviantes. Il faut d'emblée envisager une approche comparative entre des enfants tout-venant et des enfants susceptibles de présenter un développement atypique.

#### **b- MFCC, spectrogrammes, scalogrammes, mel-spectrogramme, pitch**

Afin de mener une approche comparative entre les prosodies issues d'individus sains et atteints de TSA, différentes exploitations ont été menées sur les signaux audios dans le domaine temporel comme fréquentiel. Le domaine fréquentiel permet de mettre en avant des différences de modulation des fréquences de la voix chez les enfants atteint de TSA, tandis que le domaine temporel permet de mettre en avant d'autres anomalies, comme l'allongement des voyelles.

La première exploitation sur le signal audio est la création d'un spectrogramme. Un spectrogramme est la représentation visuelle de la variation d'amplitude d'énergie des fréquences d'un signal au cours du temps. Cette représentation est obtenue après l'application d'une transformée de Fourier discrète sur des segments du signal audio préalablement segmentés, et elle permet d'identifier visuellement les formants de la voix, leurs durées, mais aussi leurs variations au cours du temps. C'est cette image qui, une fois dépourvue d'informations parasites telles que le titre, la légende ou encore les axes, sera utilisée en entrée de notre IA.

Le spectrogramme nous permet de relever beaucoup d'informations pertinentes et déterminantes, mais l'inconvénient majeur est que le spectrogramme ne tient pas compte de la perception humaine des sons. En effet, les humains ont une meilleure perception des sons de basses fréquences par rapport aux sons de haute fréquence. De manière générale, les enfants essayent de reproduire les sons qu'ils perçoivent. Ainsi, il est intéressant de mettre à l'échelle de la perception humaine le spectrogramme obtenu au préalable, notamment grâce à l'échelle Mel. L'échelle Mel est une échelle logarithmique permettant de convertir les amplitudes des fréquences réelles en amplitudes de fréquences perçues. C'est de cette manière que le Mel-spectrogramme est obtenu, et qui constituera la seconde image pour notre IA.

La troisième image extraite du signal audio est celle des MFCC. Les MFCC, ou coefficients cepstraux en fréquences de Mel, sont des caractéristiques fréquemment utilisées en reconnaissance vocale et en traitement du signal audio pour représenter les propriétés spectrales d'un audio. Ces coefficients sont

obtenus grâce à une transformée en cosinus discrète des valeurs du Mel-spectrogramme obtenus à l'étape précédente. Grâce à cette transformée, 14 coefficients numérotés de 0 à 13 sont extraits du Mel-spectrogramme. Ces 14 coefficients n'ont pas d'interprétations physiques directes, mais permettent de détecter des similitudes lors de la parole, et plus globalement de l'utilisation de la voix. Les enfants atteints de TSA rencontrant une difficulté à moduler leurs voix, des similitudes entre TSA, qui ne sont pas présentes chez les enfants sains, pourront être mises en lumière, et permettre de classifier les audios de manière efficace.

Le pitch, ou hauteur tonale, est la perception de la fréquence fondamentale chez l'Homme d'un son. La différence du comportement du pitch au cours du temps entre un enfant atteint de TSA et un enfant sain peut être anticipée. En effet, une modulation de la voix erronée chez les enfants atteints de TSA entraînera de façon systématique un comportement anormal du pitch. L'extraction du pitch est donc une caractéristique du signal audio qui peut être jugée comme pertinente à extraire.

Par ailleurs, on obtient la fréquence fondamentale en recherchant la fréquence possédant le pic d'énergie maximal, et ce, sur plusieurs segments de l'audio préalablement segmenter. On obtient donc une courbe qui traduit l'évolution de la fréquence fondamentale en fonction du temps. C'est cette évolution qui va différer selon la condition de l'enfant, étant atteint de TSA ou sain.

La dernière exploitation sur le signal audio est la création d'un scalogramme. Tout comme un spectrogramme, un scalogramme est la représentation visuelle de la variation d'amplitude d'énergie des fréquences d'un signal au cours du temps. Cependant, le scalogramme n'est pas obtenu à l'aide d'une transformée de Fourier discrète, mais grâce à une transformée en ondelettes continues. Cette différence permet d'avoir une meilleure localisation temporelle et fréquentielle des caractéristiques du signal. L'intérêt d'extraire à la fois le scalogramme et le spectrogramme est de pouvoir détecter des informations que les spectrogrammes ont atténues ou supprimées, tandis que le scalogramme les a gardées, ou inversement.

## 1.2- Filtrage de la voix

Afin d'exploiter au mieux les enregistrements des vocalises de bébés que l'on a à disposition, nous avons décidé de filtrer la voix de l'enfant. En effet, les séquences n'ayant pas été enregistrées dans des conditions propices et dédiées à un diagnostic, il y a de manière générale, beaucoup de bruits ambients, allant de bruits de couverts à la voix des parents en arrière-plan.

Ainsi, il a fallu choisir le filtre le plus adapté. Deux grands types de filtre existent : les filtres FIR et IIR. Bien que les filtres IIR puissent présenter des problèmes de stabilité et introduire une distorsion de phase, ils sont plus efficaces que les filtres FIR pour obtenir des transitions abruptes entre la bande passante et la bande atténuee. De plus, les filtres FIR nécessitent généralement un ordre plus élevé pour obtenir des performances comparables aux filtres IIR.

C'est pour cette raison que nous nous sommes dirigés vers le filtre de Butterworth, un type de filtre IIR couramment utilisé.

De surcroît, nous avions dans un premier temps opté pour un filtre passe-bande.

Les bruits ambients, tels que le bruit du vent ou le grincement des machines, ont souvent des composantes haute fréquence importantes. En les atténuant à l'aide d'un filtre passe-bas, on améliore la clarté de la parole.

Les bruits de fond, tels que le bourdonnement des moteurs ou le grondement du trafic, ont généralement des composantes basses fréquences. En les éliminant grâce à un filtre passe-haut, on réduit le bruit de fond et rend la parole plus intelligible.

Pour un bébé de 0 à 6 mois, la gamme de fréquence de sa voix se situe entre 200 et 500 Hz. Ainsi, le choix de 700 Hz pour le filtre passe-bas permet de conserver la parole tout en atténuant les bruits sifflants. Le choix de 100 Hz pour le filtre passe-haut permet d'éliminer les bruits de fond tout en préservant les graves de la voix.

Cependant, après de multiples tests, nous nous sommes aperçus qu'un filtre passe-bas était plus adapté. En effet, dans ce contexte, les bruits ambients résident dans les hautes fréquences. De plus, un filtre passe-haut altérerait la qualité du son. Malgré tout, nous avons longuement hésité quant au choix de la fréquence de coupure, car la fréquence de la voix de l'enfant se situe plus ou moins dans la même gamme que celle des bruits ambients.

Finalement, nous avons établi une fréquence de coupure à 7000 Hz, afin de ne pas détériorer la qualité du son et de garder toutes les informations. Actuellement, il reste des imperfections dans le filtrage de la voix. Néanmoins, ce sont des bruits blancs résidant sur chaque fréquence. Ils sont donc facilement distinguables et n'altèrent pas la qualité des images.

### 1.3- Génération des caractéristiques sous forme d'image

Une fois l'audio filtré, les différentes caractéristiques citées précédemment vont être extraites, comme le spectrogramme, les MFCC, etc. L'ensemble de ces caractéristiques sont extraites sous forme de représentation visuelle en diagramme de couleur ou de courbe pour le pitch.

Initialement, nous avions choisi d'extraire les caractéristiques vocales à l'aide du logiciel Matlab. Ce dernier est effectivement parfaitement adapté à l'analyse de données et à l'extraction d'images. Les fonctions disponibles permettent ainsi de visualiser les courbes et les diagrammes dont nous avions besoin. Cette première approche nous a permis de récupérer toutes les caractéristiques correspondant à chaque piste vocale de la base de données, grâce à un programme automatisé. Ainsi, nous avons pu faire un travail en amont d'observation et de comparaison des aspects discriminants des images, détaillées dans la suite du rapport.

Toutefois, bien que ce logiciel était le plus pertinent et le plus adapté, nous avons dû convertir ce programme pour l'utiliser avec Python. En outre, le reste du travail à effectuer était plus accessible sur ce langage, notamment pour l'IA et pour l'application. Ainsi, pour des questions de cohérence et de compatibilité, nous avons dû adapter le travail déjà effectué.

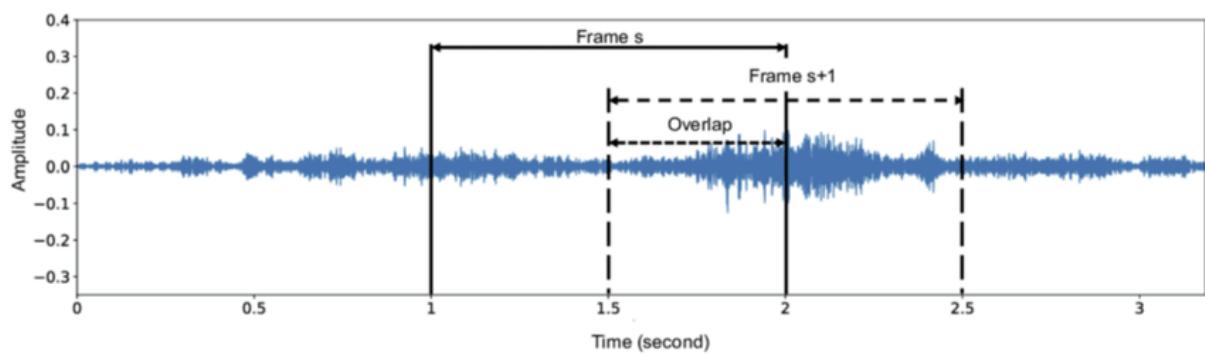
Par conséquent, le programme a été développé à l'aide de la bibliothèque `scipy` et `librosa`. Ces bibliothèques permettent d'effectuer des traitements de signaux poussés et de manière efficace, c'est pourquoi elles sont utilisées à tous les stades de l'extraction des caractéristiques vocales, c'est-à-dire lors des transformations effectuées sur l'audio filtré, mais aussi lors de la création finale du diagramme ou de la courbe.

Pour la création des spectrogrammes, des mel-spectrogrammes et des MFCC, les signaux audios filtrés sous forme de tableau de valeur sont utilisés en argument des fonctions `scipy` ou `librosa`.

associées (`spectrogramme()` de `scipy` pour les spectrogrammes, `pyin()` de `librosa` pour le pitch, `melspectrogramme()` de `librosa` pour les mel-spectrogrammes, et `mfcc()` de `librosa` pour les mfccs).

Cependant, pour la totalité de ces caractéristiques vocales, les méthodes d'extraction incluent une étape de segmentation de l'audio en segments de quelques millisecondes. Une caractéristique de ces segments, nécessaire pour assurer l'exactitude des résultats de ces méthodes, est la stationnarité des segments. Autrement dit, le signal est le même sur la période de temps. Pour que les segments soient considérés comme stationnaires, leur durée est de 30 millisecondes, soit la durée pour laquelle la voix humaine est considérée comme stationnaire.

Toutefois, cette segmentation peut aussi apporter une perte de données, notamment des données entre 2 segments adjacents. Pour éviter cette perte, les segments sont superposés de 25 millisecondes. Les segments sont dits superposés quand le début de chaque segment se situe à l'intérieur du segment précédent, comme l'image ci-jointe le montre.



Bien que cette superposition supprime la perte de données, elle amplifie également les données situées dans la période de superposition, ce qui pourrait amener des biais et des incohérences ultérieurement. C'est pourquoi, en plus de la superposition, chaque segment est modifié. En effet, chaque segment est multiplié par une fenêtre. Dans notre cas, une fenêtre de Hamming, générée par `scipy`, multiplication qui a pour conséquence d'annuler l'amplification injustifiée de certaines sections de l'audio, due à la segmentation superposée de ce dernier. Ce processus s'appelle le Windowing, ou fenêtrage en français.

Les données obtenues à l'aide de ces fonctions sont ensuite transformées en image à l'aide de 3 fonctions qui sont : `plot` de `matplotlib` pour créer la courbe du pitch, `specshow` de `librosa` pour les diagrammes des MFCC ainsi que des mel-spectrogrammes, et enfin `pcolormesh` de `matplotlib` pour les spectrogrammes, permettant des paramètres plus personnalisés, tels que la couleur ou encore le dégradé, car les besoins sur les spectrogrammes sont plus élevés.

Les images obtenues vont être interprétées par les intelligences artificielles en tant que tableau de valeurs en 3 dimensions, qui sont la hauteur, la largeur et une dernière dimension qui représente les 3 canaux de couleurs RGB. Pour faciliter le traitement futur de ces images par une intelligence artificielle, ces images sont dépourvues de légende, de titre ou encore d'axe pour éviter toutes informations superflues non discriminantes entre les enfants sains et atteints de TSA. Le choix des couleurs des diagrammes a aussi une grande importance pour augmenter la précision des intelligences artificielles. En effet, les diagrammes doivent avoir des couleurs sur la plus grande portion possible du spectre des couleurs visibles, et ce, pour que les 3 canaux RGB soient utilisés par l'intelligence artificielle, et donc que ces 3 canaux apportent une information déterminante et non aléatoire.

En effet, si un diagramme n'est composé que de nuances de rouge, les canaux liés aux couleurs verte et bleue n'apporteront aucune information, ce qui aura pour conséquence de compacter l'information sur un seul canal, le canal rouge, entraînant des difficultés et une perte de précision sur les différents modèles d'intelligence artificielle. Pour ce faire, des gammes de couleurs similaires aux gammes de couleurs type « arc-en-ciel » sont utilisées pour répartir un maximum l'information sur l'ensemble des 3 canaux.

#### 1.4- Etude comparative

Après avoir obtenu toutes les images, nous avons effectué des observations préliminaires afin de visualiser les caractéristiques vocales qui changent d'un enfant atteint de TSA à un enfant sain.

Dans un premier temps, nous avons réalisé un tableau regroupant toutes les caractéristiques (MFCC, spectrogrammes, pitchs, scalogrammes et mel-spectrogrammes) en fonction de la condition de l'enfant (sain ou TSA), de son sexe, de son âge et de son activité.

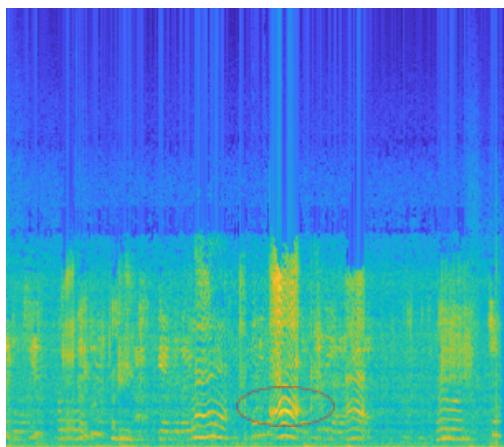
Grâce à ce tableau, nous avons pu effectuer une première analyse des images et constater des différences significatives à l'œil nu.

Cela nous a également permis de savoir quelles caractéristiques étaient pertinentes et exploitables. En particulier, nous avons réalisé que les pitchs et les scalogrammes n'étaient pas significatifs. C'est pourquoi nous avons décidé de les mettre de côté.

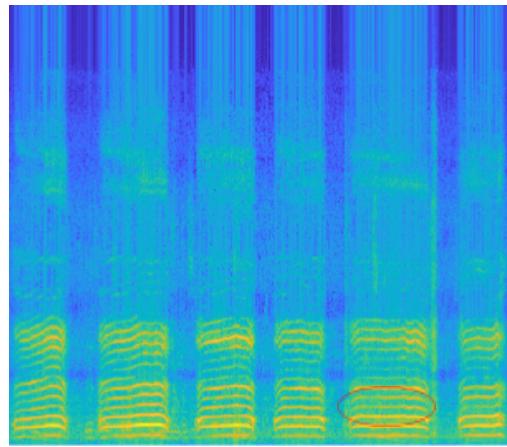
Concernant les pitchs, nous ne les avons pas exploités vu que leur allure n'était pas assez nette, donc l'IA n'aurait pas su en tirer de conclusion.

Au sujet des spectrogrammes, nous avons remarqué que, sur ceux des enfants sains, il y a autant de contours descendants que de contours montants alors que sur ceux des enfants atteints de TSA, les contours sont plus lisses : il n'y a pas de contours, ni montants, ni descendants.

Spectrogramme d'un enfant sain :

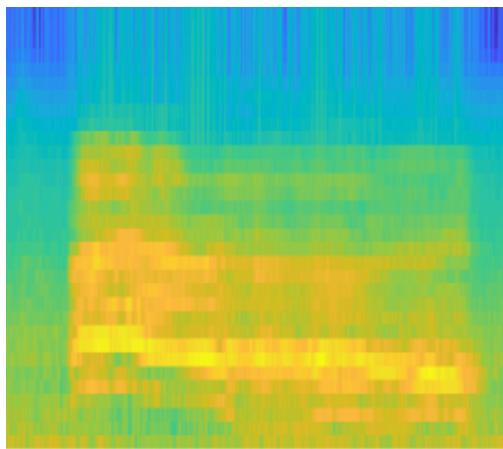


Enfant atteint de TSA :

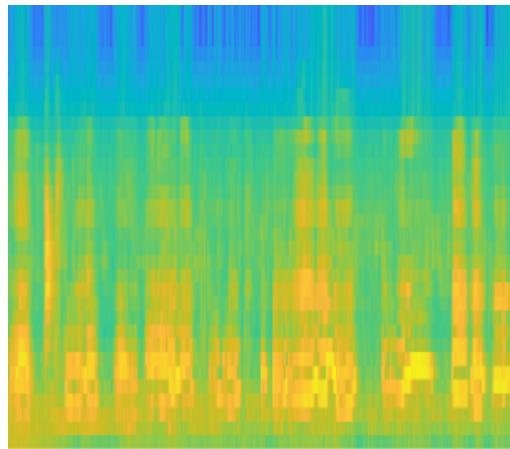


À l'égard des Mel-spectrogrammes, nous remarquons que ceux des enfants sains sont réguliers avec des bandes de fréquences fondamentales fines et similaires; or, sur ceux des enfants atteints de TSA, on observe des bandes allongées et irrégulières.

Mel- spectrogramme enfant sain :

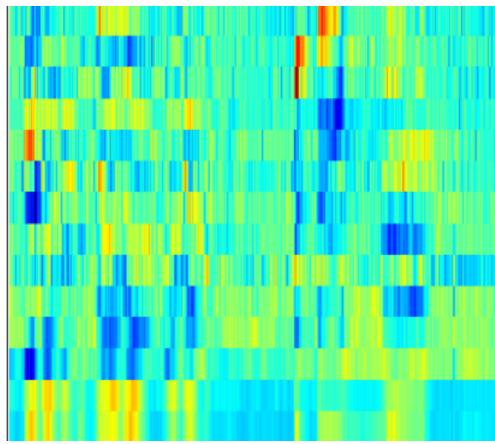


Enfant atteint de TSA :

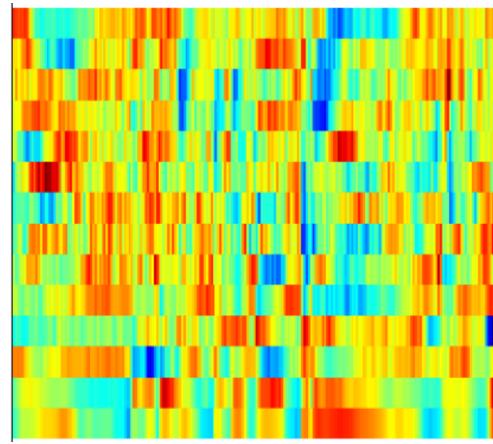


En ce qui concerne les MFCC, ceux des enfants sains sont globalement plus homogènes (en termes de couleurs correspondant à la puissance des coefficients) alors que sur ceux des enfants atteints de TSA il y a une prédominance des couleurs rouge et jaune, ce qui reflète une plus grande puissance des coefficients.

MFCC enfant sain :



Enfant atteint de TSA :

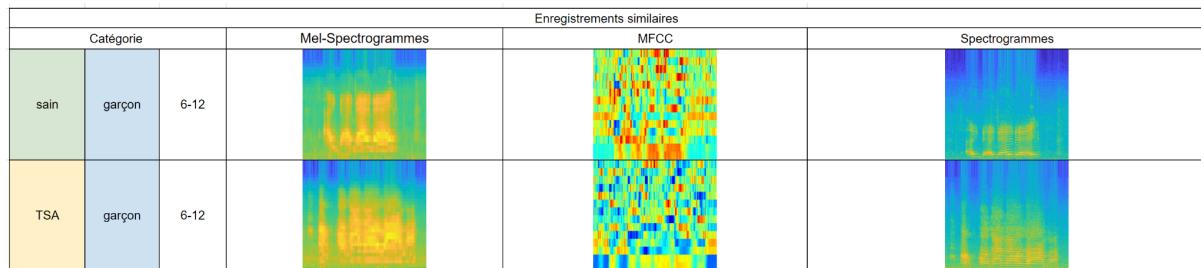


Par la suite, nous avons constaté que des comparaisons avec des enregistrements audio de mêmes durées sont plus pertinentes et rigoureuses. Effectivement, dans nos premières comparaisons, nous n'avions pas nécessairement les mêmes échelles, ce qui pouvait biaiser nos analyses surtout en ce qui concerne les analyses des mel-spectrogrammes. Ainsi, nous avons choisi de réaliser les mêmes comparaisons, mais avec des audios de même durée, soit 10 secondes, afin d'obtenir la même échelle pour nos analyses. [cf. annexe 2]

Pour ce faire, nous avons dupliqué notre base de données initiale pour générer une nouvelle base de données à rogner. Les enregistrements longs ont été coupés en séquences audios de 10 secondes. Les séquences qui faisaient moins de 8 secondes ont été ensuite supprimées de notre nouvelle base de données.

Pour les mêmes raisons, ces modifications étaient également plus pertinentes dans l'utilisation de l'IA.

Après ces deux différentes analyses, nous avons constaté que les comparaisons réalisées n'étaient pas les plus appropriées pour analyser de manière approfondie les différentes caractéristiques. Notre choix s'est donc porté sur la comparaison de deux enregistrements de mêmes durées et similaires à l'écoute. Ils ont donc été sélectionnés pour chaque condition, l'un d'un enfant atteint de TSA et l'autre d'un enfant sain. Toutefois, dans chaque cas, il concerne un garçon âgé de 6 à 12 mois avec une activité non spécifiée.



Sur ce tableau, on remarque que sur le spectrogramme de l'enfant sain, les contours sont symétriquement montants et descendants. Or, chez l'enfant atteint de TSA, les contours n'ont pas de variations et sont très plats. Ce manque de variations peut être expliqué par le manque d'intonations dans les vocalises d'un enfant atteint de TSA. De plus, les enfants sains marquent des pauses nettes, contrairement aux enfants atteints de TSA.

Sur les mel-spectrographmes, on voit que la plage de variation en fréquence fondamentale de l'enfant TSA est plus étendue que celle de l'enfant sain, qui est très étroite. L'enfant atteint de TSA parle également à des fréquences plus hautes que l'enfant sain.

En ce qui concerne les caractéristiques différentes des MFCC, nous ne sommes pas parvenus à tirer de conclusion pertinente du fait que nos observations précédentes n'étaient plus valables dans ce contexte.

## Partie 2 : L'IA pour l'aide au diagnostic

### 2.1- Choix de l'architecture de l'IA

Dans le cadre de l'étude de nos signaux, nous avons employé des modèles d'intelligence artificielle afin d'améliorer la précision et l'efficacité de nos analyses. Ces modèles avancés nous ont permis de détecter et d'interpréter des motifs complexes dans les données spectrales, facilitant ainsi une compréhension plus approfondie des phénomènes observés.

L'objectif final étant de pouvoir interpréter un signal extérieur à l'un de notre base de données, il était donc nécessaire de construire un modèle avec une bonne architecture qui puisse fournir une précision suffisante.

Il existe ainsi deux modèles d'architecture utilisables : un auto-encodeur ou un réseau de neurones pré-entraîné via la méthode de transfer learning.

Nous avons ainsi étudié les deux méthodes à la fois afin de déterminer la plus efficace.

En premier lieu, les auto-encodeurs produisent en sortie une reconstruction de l'entrée. En outre, un auto-encoder est constitué de deux sous-réseaux : un encodeur et un décodeur. Lors de l'apprentissage, l'encodeur apprend un ensemble de caractéristiques à partir des données d'entrées qu'on appelle la représentation latente. En parallèle, le décodeur est entraîné à reconstruire les données à partir de ces caractéristiques. L'auto-encoder peut alors être appliqué afin de prédire des entrées qui n'ont pas été précédemment observées.

Les auto-encoders sont très faciles à généraliser et peuvent être utilisés sur différents types de données, notamment des images, des séries temporelles ou du texte.

Ainsi, nous nous sommes naturellement tournés vers l'implémentation d'un modèle auto-encoder sur Matlab, puis sur Python.

Néanmoins, en raison de l'hétérogénéité de notre base de données, nous ne sommes pas parvenus à implémenter un auto-encoder. Compte tenu du fait qu'un auto-encoder, lors de son apprentissage, apprend un schéma de compression-décompression, notre base de données ne permettait pas suffisamment à l'auto-encoder de différencier des spectrogrammes 'Sains' et 'TSA'. C'est pourquoi lors des tests, même sur un petit jeu de données contenant uniquement des spectrogrammes issus du dossier 'TSA', l'auto-encoder les classait en tant que spectrogrammes 'Sains'.

La perspective d'utilisation d'un auto-encoder pour nos données a donc finalement été écartée.

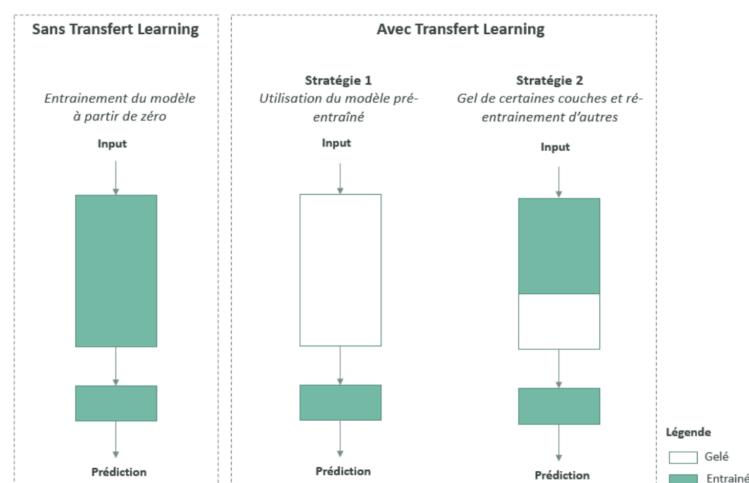
En second lieu, nous nous sommes donc tournés vers le Transfert Learning.

C'est une méthode de Machine Learning qui repose sur l'utilisation de modèles préalablement entraînés sur un grand nombre de données. Le Transfert Learning est composé de deux étapes principales. La première, qui est le pré-entraînement des données, permet au modèle d'apprendre des caractéristiques générales. Cependant, cette étape est déjà effectuée lorsque nous récupérons le modèle.

La deuxième étape consiste à ajuster (ou "fine-tuner") le modèle sur le jeu de données spécifique à la tâche souhaitée, dans notre cas à des spectrogrammes.

Le Transfert Learning est particulièrement flexible et peut être appliquée à divers types de données, y compris des images, du texte et des données audio.

Ainsi, nous nous sommes naturellement tournés vers l'implémentation d'un modèle de Transfert Learning, d'abord sur Matlab, puis sur Python.



Représentation du Transfer Learning

## 2.2- Etapes préliminaires

### a- Présélection des réseaux de neurones pré-entraînés

La première étape pour implémenter le Transfert Learning à notre projet a été de choisir un réseau de neurones pré-entraînés. Pour cela, nous nous sommes appuyés sur un tableau ainsi que sur un schéma comparatif des IA disponibles sur Matlab.

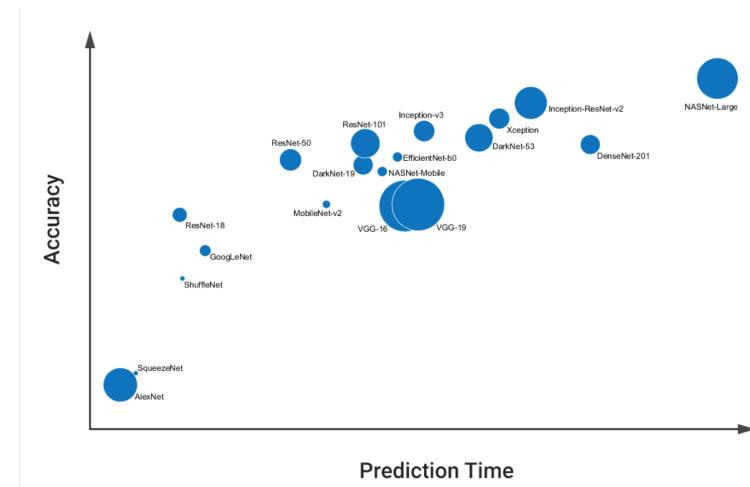


Tableau comparatif des modèles disponibles sur Matlab

Puis, nous avons choisi de faire une présélection des modèles qui nous semblaient les plus pertinents afin de pouvoir tester un maximum de modèles et finalement choisir celui ou ceux avec les meilleurs résultats.

Le premier réseau que nous avons choisi de prendre est **AlexNet**, avec seulement 25 couches. C'est le réseau avec l'architecture la plus simple et le plus rapide à exécuter, nous permettant ainsi de mieux appréhender les modèles pré-entraînés. Il constitue une première approche, en utilisant le plus simple à comprendre.

Par la suite, les critères de choix pour les autres modèles ont été plus difficiles à établir, mais nous avons pu isoler 3 modèles : MobileNet, VGG-16 et ResNet-50.

Tout d'abord, notre choix s'est porté sur **MobileNet**, un modèle avec une architecture un peu plus complexe, mais optimisée, qui permet donc d'avoir une bien meilleure précision qu'avec AlexNet, pour un temps d'exécution légèrement plus élevé.

Par la suite, nous avons choisi d'utiliser **VGG-16**, qui est un modèle avec une architecture plus profonde, ce qui permet au réseau de capturer des caractéristiques plus complexes mais nécessite une puissance de calcul beaucoup plus élevée. En effet, à titre de comparaison, VGG-16 contient 138 Millions de paramètres contre 3,5 Millions pour MobileNet. Ce réseau nous permet alors de voir si un modèle très complexe nous permettrait d'avoir de meilleurs résultats.

Le dernier réseau sélectionné a été **ResNet-50**, qui est assez équilibré entre profondeur et performance

avec 25.6 Millions de paramètres. Ce réseau très populaire possède une méthode d'apprentissage assez particulière en utilisant des blocs résiduels.

Ces blocs résiduels aident les réseaux de neurones à apprendre plus efficacement. Ils utilisent des "raccourcis" qui permettent aux informations de circuler plus facilement à travers le réseau, même s'il est très profond. Ces raccourcis aident à éviter les problèmes qui rendent l'apprentissage difficile dans les réseaux profonds, tels que les problèmes de dégradation des gradients, ce qui améliore les performances et la précision du modèle.

## b- Prétraitement des données

Une fois les réseaux choisis, il y a eu plusieurs étapes de prétraitement des données importantes avant d'entamer les tests.

Tout d'abord, il a fallu adapter les dimensions de nos données en fonction de celles prises en compte par les modèles d'IA sélectionnés. Les spectrogrammes générés par Matlab avaient des dimensions de 607 x 536, qui ne sont pas acceptées par les modèles. Nous les avons donc redimensionnés en 224 x 224.

Concernant le format, nous avons décidé de les laisser en PNG, car c'est un format qui n'a pas de perte de qualité. De plus, avec une base de données plutôt réduite, cela ne posait pas de problèmes de stockage.

Nous nous sommes ensuite questionnés sur les couleurs des spectrogrammes générés par Matlab. Ceux par défaut étaient dans une gamme de vert, cependant les modèles d'IA comparent les valeurs RGB. Ainsi, avec le rouge et le bleu à 0, nous réduisons les capacités du modèle, c'est pourquoi nous nous sommes orientés vers des spectrogrammes utilisant toutes les gammes RGB.

Deux autres étapes souvent utilisées à des fins de classification dans le prétraitement de données sont la normalisation de données et l'augmentation de données.

La normalisation consiste à diviser chaque pixel par 255 pour qu'ils soient compris entre 0 et 1 plutôt que 0 et 255, ce qui est plus simple à traiter pour le réseau. Cependant, après plusieurs tests, nous nous sommes aperçus que nous obtenions de meilleurs résultats avec les données non normalisées.

```
train_data_dir = "/Users/nathanlecoin/Desktop/VSCODE/jet/train"
validation_data_dir = "/Users/nathanlecoin/Desktop/VSCODE/jet/validation"

# Paramètres
img_width, img_height = 224, 224
batch_size = 32
epochs = 30

# Préparer les générateurs de données
train_datagen = ImageDataGenerator(preprocessing_function=keras.applications.vgg16.preprocess_input)
val_datagen = ImageDataGenerator(preprocessing_function=keras.applications.vgg16.preprocess_input)

train_generator = train_datagen.flow_from_directory(
    train_data_dir, target_size=(img_width, img_height), batch_size=batch_size, class_mode="binary"
)

validation_generator = val_datagen.flow_from_directory(
    validation_data_dir, target_size=(img_width, img_height), batch_size=batch_size, class_mode="binary"
)
```

Prétraitement des données

L'augmentation, quant à elle, consiste à effectuer des légers déplacements sur le côté ou à retourner les images pour que le modèle s'entraîne sur différentes positions et soit plus performant sur des nouveaux cas jamais traités. Toutefois, cette solution n'est pas très pertinente dans le cas de spectrogrammes, car ils se lisent uniquement de gauche à droite et ils perdent beaucoup d'informations après cette étape.

Pour ces raisons, nous avons décidé de ne pas les implémenter dans nos modèles d'IA.

### 2.3- Tests et entraînements

#### a- Réalisation des tests sur Matlab

Une fois les étapes de prétraitement terminées, nous avons dû choisir le langage à utiliser. Sachant que nos images sont générées par Matlab et que ce logiciel propose un large choix de réseaux pré-entraînés pour des images, cela nous semblait être la meilleure solution.

Nous avons donc commencé en utilisant le réseau AlexNet, comme évoqué précédemment, afin d'obtenir des premiers résultats le plus rapidement possible.

Après avoir séparé notre jeu de données en données d'entraînement et de validation, avec une répartition de 80/20 respectivement, et mis en place des paramètres adéquats à l'entraînement de l'IA sur notre jeu de données, nous avons réussi à obtenir des résultats allant jusqu'à 80% de précision. Ce premier résultat nous paraissait correct, mais encore loin de ce que nous projetions d'avoir.

Nous avons donc par la suite changé le réseau de neurones pré-entraînés en prenant le réseau ResNet-50. Cependant, ce réseau ayant jusqu'à sept fois plus de paramètres que AlexNet, nous avions besoin d'une plus grande puissance de calcul pour entraîner notre modèle.

En effet, un réseau de neurones artificiel est un ensemble de calculs mathématiques complexes. Il est donc nécessaire de faire tourner un réseau de neurones sur une machine puissante, avec beaucoup de ressources de calculs, d'autant plus lorsque c'est un réseau assez profond. Pour ce faire, la solution la plus efficace est l'utilisation de processeurs graphiques (GPU). Un GPU est un circuit électrique capable d'effectuer des calculs mathématiques à grande vitesse. Les tâches informatiques telles que le rendu graphique, le Machine Learning et le montage vidéo nécessitent l'application d'opérations mathématiques similaires sur un grand jeu de données. Néanmoins, avec Matlab, l'application n'accepte qu'un type de GPU très spécifique qui n'était présent ni sur les ordinateurs mis à notre disposition dans l'école, ni sur nos ordinateurs personnels respectifs.

Or, sans GPU, le logiciel n'était plus capable de compiler les modèles d'IA.

D'autre part, en parallèle de cette étape, nous avons commencé à développer l'application sur Python, et l'utilisation d'IA sur Matlab rendait l'exploitation beaucoup plus compliquée.

C'est pourquoi nous avons décidé de convertir le code Matlab afin de l'exploiter sur Python. Dès lors, nous avons continué à travailler intégralement sur Python, dans l'optique d'optimiser l'utilisation de l'IA, mais également de faciliter le développement futur de l'application.

## b- Réalisation des tests sur Python

Ainsi, nous avons basculé sur le Transfer Learning à l'aide de Python, notamment en utilisant l'outil "open source" TensorFlow. Pour cela, nous avons utilisé l'éditeur de code Visual Studio Code et installé Python ainsi que toutes les librairies nécessaires pour faire fonctionner nos réseaux de neurones.

Tout compte fait, nous avons décidé d'implémenter deux réseaux : VGG16 et MobileNet. AlexNet a été écarté, car il n'était pas pris en charge par la librairie Keras. Quant à ResNet50, ce dernier causait certains problèmes dans l'application et dans la mise en cascade de nos réseaux.

```
base_model = VGG16(weights="imagenet", include_top=False, input_shape=(img_width, img_height, 3))

x = base_model.output
x = GlobalAveragePooling2D()(x)
x = Dense(1024, activation="relu")(x)
predictions = Dense(1, activation="sigmoid")(x)

model = Model(inputs=base_model.input, outputs=predictions)

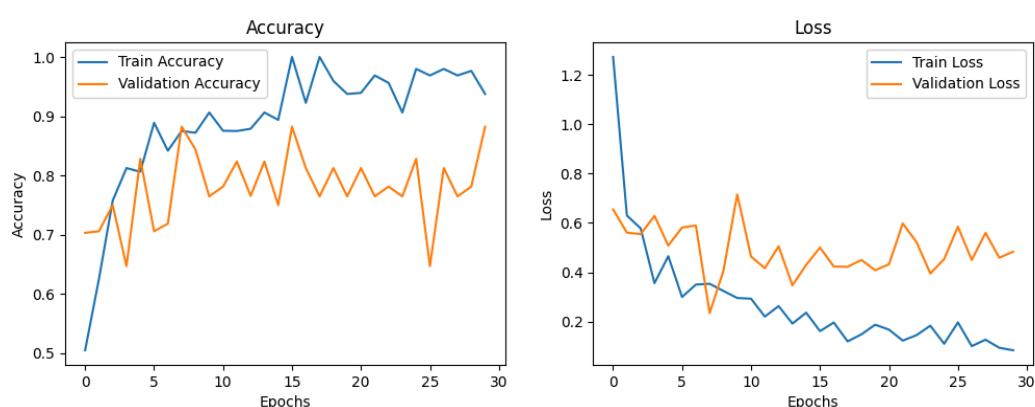
for layer in base_model.layers:
    layer.trainable = False

model.compile(optimizer="adam", loss="binary_crossentropy", metrics=["accuracy"])
model.summary()
```

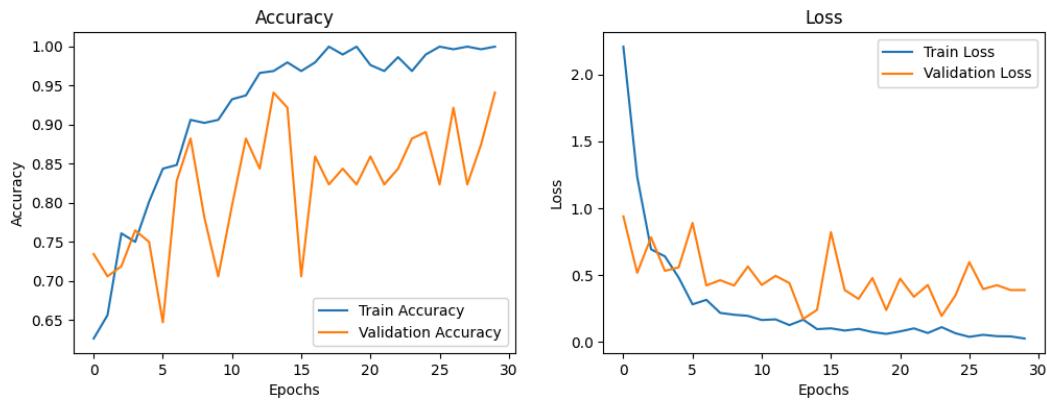
Implémentation de VGG16

Afin de mettre en place chaque réseau, nous avons procédé de la même manière pour l'un et l'autre. Tout d'abord, nous avons séparé notre base de données en deux sous-ensembles : 'train' et 'validation', avec une répartition de 80% et 20%. Les réseaux s'entraînent ainsi sur la grande partie des données. Puis, les données 'validation' servent à vérifier si les réseaux n'effectuent pas ce qu'on appelle le 'surapprentissage'.

À la fin des tests d'entraînement, nous obtenons les résultats suivants :



Résultats d'entraînement et de validation de MobileNet



Résultats d'entraînement et de validation de VGG16

```

/Users/nathanlecoin/Desktop/VSCode/jet/validation/TSA/Copie_de_MG1a_jeune_spectrogramme.png: TSA
/Users/nathanlecoin/Desktop/VSCode/jet/validation/TSA/Copie_de_MG1b_segment1_spectrogramme.png: TSA
/Users/nathanlecoin/Desktop/VSCode/jet/validation/TSA/Copie_de_MG22a_2ans_joue_spectrogramme.png: TSA
/Users/nathanlecoin/Desktop/VSCode/jet/validation/TSA/Copie_de_MG21_1an_jouepleure_spectrogramme.png: TSA
/Users/nathanlecoin/Desktop/VSCode/jet/validation/TSA/Copie_de_MG7c_stimule_segment1_spectrogramme.png: TSA
/Users/nathanlecoin/Desktop/VSCode/jet/validation/TSA/Copie_de_MF2a_segment1_spectrogramme.png: TSA
/Users/nathanlecoin/Desktop/VSCode/jet/validation/TSA/Copie_de_MG23a_18mois_pleur_spectrogramme.png: TSA

Total predictions: 35
Number of TSA predictions: 31
Number of Sain predictions: 4

```

Prédictions de VGG16 sur un ensemble de 35 spectrogrammes TSA

Par ailleurs, à l'issue de cette étude, nous avons choisi de ne pas étudier les Mel-spectrographes ainsi que les MFCC et le pitch affiné. Les tests des IA présentant de trop mauvais résultats, même en essayant différentes palettes de couleurs pour mettre en lumière certaines données essentielles à l'apprentissage, nous avons jugé préférable de ne pas les intégrer au diagnostic.

L'étape finale de la partie sur les IA réside dans la sauvegarde des réseaux pré-entraînés sur nos spectrogrammes, en fichier HDF5 (Hierarchical Data Format version 5). C'est une extension utilisée pour stocker et organiser de grandes quantités de données. Finalement, ces fichiers alimenteront l'application pour l'aide au diagnostic.

## Partie 3 : Développement de l'application d'aide au diagnostic

Afin de rendre des scripts Python accessibles et faciles d'utilisation, la création d'une application simpliste est une approche pratique.

Dans un premier temps, nous avons dû nous pencher sur la question des systèmes d'exploitation sur lesquels notre logiciel sera disponible, comme les appareils mobiles, les ordinateurs de bureau, ou d'autres plateformes. Initialement, nous souhaitions le mettre en place sur smartphone. Cependant, par réalisme et en raison de l'accompagnement psychologique qu'il nécessite, nous avons décidé de renoncer à cette option.

En effet, nous ne souhaitions pas que l'application soit grand public, car si le logiciel n'est pas correctement utilisé et interprété, trop de familles seraient informées par défaut, leur causant du stress supplémentaire. De plus, le risque serait aussi de méprendre les TSA pour une autre condition, comme des retards de langage ou des troubles de l'audition et du neurodéveloppement, s'il n'y a pas de corps médical pour encadrer l'utilisation.

L'application doit alors être un logiciel d'outil médical. Par conséquent, il serait uniquement utilisé dans le cadre de l'aide au diagnostic par des médecins agrés.

De surcroît, un smartphone nous aurait limités par son processeur, moins puissant qu'un ordinateur.

Nous nous sommes alors orientés vers la création d'un logiciel sur ordinateur, notamment sur Mac et Windows, avec le .app et le .exe.

Ainsi, plusieurs solutions s'offraient à nous dans le choix du langage de développement. Toutefois, au vu du temps que l'on disposait et de la faisabilité, nous avons décidé de coder l'entièreté de l'application en Python et d'utiliser la bibliothèque graphique libre d'accès, Tkinter.

Il a alors fallu structurer notre projet de façon à ce que le programme soit efficace et puisse tourner sur n'importe quelle machine.

Pour cela, nous avons dû revoir l'entièreté des codes déjà produits, que ce soit pour générer des figures ou bien pour les étudier avec nos modèles d'IA. L'objectif vers lequel nous nous sommes dirigés était que tout soit automatisé. L'utilisateur n'aurait donc rien de technique à réaliser.

On a alors réalisé notre code en trois programmes :

Le premier est '**generate.py**', le programme utilisé pour générer des figures à partir d'un audio. Ce code fut optimisé afin de pouvoir rendre chaque tâche indépendante l'une de l'autre. Nous avons fait le choix de créer des fonctions de lecture de fichiers audio et une pour le filtrer :

```
def mic():
    fs = 44100 # Fréquence d'échantillonnage
    duration = 10 # Durée de l'enregistrement en secondes
    myrecording = sd.rec(int(duration * fs), samplerate=fs, channels=1)
    sd.wait() # Attendez la fin de l'enregistrement
    sf.write("record.wav", myrecording, fs)
    file_path = "record.wav"
    return run(file_path, get_next_directory())
```

Cette fonction permet, à partir du micro de l'ordinateur, d'enregistrer les dix premières secondes dans un format wav et de sauvegarder l'audio au nom de ‘record.wav’ dans le dossier courant. Elle utilise la librairie sounddevice qui contient la fonction rec().

```
def nomic():
    root = tk.Tk()
    root.withdraw()
    file_path = tk.filedialog.askopenfilename(filetypes=[("WAV files", "*.wav")])
    root.destroy()
    if not file_path:
        print("Aucun fichier sélectionné")
        exit()
    return run(file_path, get_next_directory())
```

Quant à la fonction nomic(), elle permet d'ouvrir une fenêtre de dialogue pour sélectionner l'audio à analyser (.wav).

Ces deux méthodes appellent la fonction run(file\_path, get\_next\_directory()) avec file\_path le chemin d'accès de l'audio et get\_next\_directory() la fonction qui retourne le nom du prochain dossier résultat.

```
def run(file_path, output_dir):
    if not os.path.exists(output_dir):
        os.makedirs(output_dir)

    audio, sample_rate = librosa.load(file_path, sr=None)
    filtered_audio = filter_audio(audio, sample_rate)

    name = os.path.splitext(os.path.basename(file_path))[0]

    window_length = round(0.03 * sample_rate)
    overlap_length = round(0.025 * sample_rate)
    return filtered_audio, sample_rate, name, output_dir, window_length, overlap_length
```

La fonction run va alors créer ce dossier résultat et retourner certaines valeurs comme la longueur de la fenêtre, l'overlap, le nom du fichier audio et, plus important, l'audio filtré par la fonction filter\_audio().

Toujours au sein de ce programme, on retrouve alors les fonctions de calcul, expliquées précédemment (do\_voice\_waveform, do\_spectrogram, do\_pitch, do\_mel\_spectrogram et do\_mfcc).

Toutes ces méthodes renvoient vers la fonction save\_fig qui enregistre la figure produite dans le même dossier.

```
save_fig(data, sample_rate, name, output_dir, suffix)
```

Le second programme ‘ia.py’ n'a pas subi de modification pour l'implémentation sur le logiciel. On se retrouve donc avec les fonctions :

- predict\_res(path), permettant de démarrer la prédiction sur l'image qui se trouve au chemin d'accès ‘path’
- load\_models(), lancée au début du programme, servant à charger les modèles pré-entraînés

Le troisième est ‘app.py’, le programme principal qui lance l'application.

Il est structuré de telle sorte qu'à chaque démarrage, il crée une instance de Tkinter permettant d'afficher une fenêtre avec certains labels et boutons qui s'afficheront et disparaîtront lorsque certaines fonctions seront appelées.

Nous avons essayé de rendre l'interface simple avec une barre de menu regroupant un onglet ‘options’ pour aller sur la page principale ou quitter l'application, un autre afin de lancer un ‘diagnostic’ (qu'il soit à partir d'un enregistrement direct ou à partir d'un fichier audio déjà enregistré), du menu ‘résultats’ dans le but de consulter le résultat d'un ancien patient et d'un menu ‘informations’.

C'est à partir de chacun de ces menus que des commandes seront exécutées. Certaines d'entre elles font appel à des fonctions, qu'elles soient du programme ‘app’, de ‘generate’ ou bien de ‘ia’.

Après avoir détaillé la structure du projet, nous allons donc expliquer comment, à partir de l'interface graphique, un résultat peut s'afficher.

Lorsque le programme app.py est lancé, l'interface graphique s'affiche avec la barre de menu et la page principale. Ceci est causé par l'initialisation des labels, boutons, frames et scrollbar, window = Tk() qui crée une fenêtre et window.mainloop() qui crée la boucle principale de l'application. De plus, l'appel des fonctions menu\_bar() et main\_menu() est réalisé dans le main().

```
def main_menu():
    """
    Affiche le menu principal de l'application.

    Elle masque tous les widgets et affiche le logo, le texte et l'étiquette du médecin à l'écran.
    """

    hide_widgets()
    logo_label.pack(pady=10)
    description_label.pack()
    doctor_label.pack(pady=10)
```

Ainsi, à chaque démarrage de l'application, les widgets sont cachés (via une fonction hide\_widgets permettant de pack\_forget() chaque élément existant), et le logo de prosod'IA est affiché ainsi que sa description et une autre photo.

Pour plus de clarté, voici à quoi ressemble l'initialisation de ‘description\_label’ :

```
description_label = Label(
    window,
    text="L'application d'aide au diagnostic précoce des TSA",
    font=("Roboto Serif", 15),
    bg=background_color,
    fg=font_color,
)
```

Elle est accessible ici, car chaque widget est défini globalement, c'est-à-dire qui peut être accessible à toutes les fonctions et classes du script.

Dès lors, l'utilisateur peut lancer une analyse en cliquant sur le menu diagnostic puis le sous-menu “importer” par exemple. Ce menu va ainsi lancer la commande menu\_import de ‘app.py’.

```
enregistrement_menu.add_command(label="Importer audio", command=menu_import)
```

Cette commande va alors supprimer les widgets qui étaient présents et en afficher de nouveaux comme un bouton d'importation :

```
import_button.pack(pady=20)
```

Si l'utilisateur clique alors sur le bouton, une nouvelle commande sera exécutée, ici run car c'est la commande associée au 'import\_button' :

```
import_button = Button(  
    window,  
    text="Importer mon audio (format .wav)",  
    font=("Roboto Serif", 15),  
    command=run,  
    bg=background_color,  
    fg=font_color,  
)
```

La fonction run est alors appelée. Tout d'abord, elle masque certains widgets de l'interface graphique et en affiche d'autres. Puis, elle appelle la fonction no\_mic() de 'generate.py' et stocke les variables retournées.

De plus, elle appelle la fonction generate.do\_voice\_waveform() et affiche la figure produite, alors que la fonction generate.do\_spectrogramme() est exécutée grâce à l'utilisation de threads. Cela a pour effet de gagner en temps et de permettre de régler des soucis d'exécution de fonctions dites asynchrones (fonctions longues dont nous n'attendons pas de résultats).

Enfin, une fois que le thread qui exécutait do\_spectrogramme est fini, c'est-à-dire lorsque le spectrogramme de l'audio importé est généré, la fonction predict\_res(path) est appelée à son tour, avec comme paramètre path, le chemin d'accès du fichier produit par do\_spectrogramme(). On récupère alors sa valeur de retour, qui est le texte contenant le pourcentage de suspicion, et on l'affiche à l'écran :

```
result, path_fig = ia.predict_res(generate.get_result_directory())  
analyse_label.pack_forget()  
result_label.config(text=result)  
result_label.pack()
```

Il en va de même pour chaque fonctionnalité du programme. Il est possible de retrouver le projet dans son intégralité, avec les codes commentés et mis à jour, grâce à la commande :

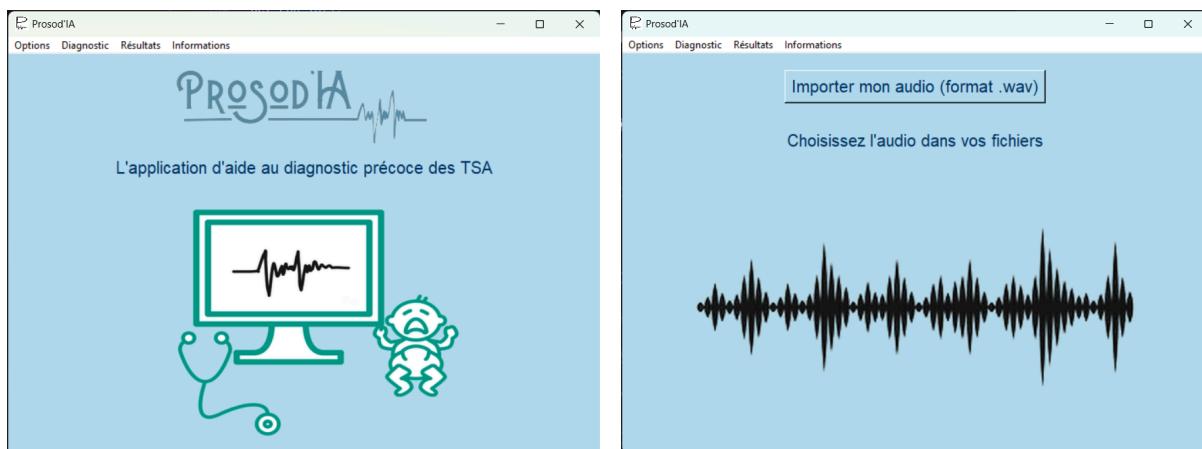
```
git clone https://oauth2:gpat-M_W-m6x1f_2XVq4VtrwJ@git.esiee.fr/lalieur/prosod-ia
```

Pour le déploiement de l'application en exécutable sur Windows et Mac, c'est grâce à py2app et auto-py-to-exe, une interface graphique simple basée sur PyInstaller, que nous avons pu mettre notre système en place.

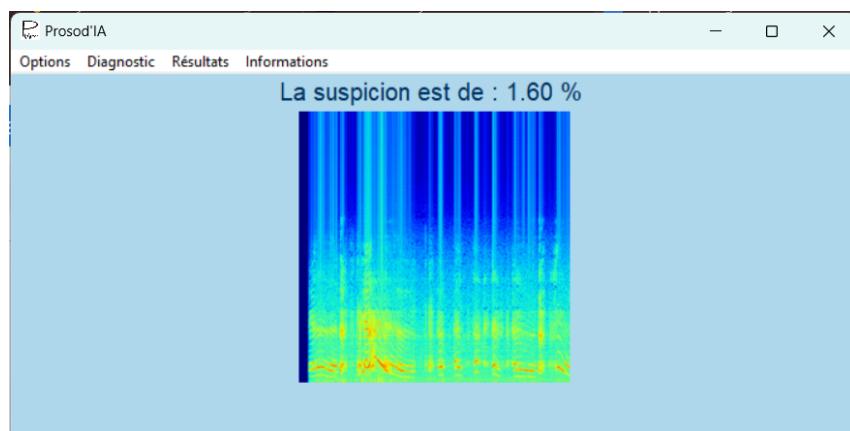
Ainsi, on obtient un exécutable qui a l'avantage d'être isolé. Ce fichier exécutable est autonome et n'a pas besoin d'une installation Python en amont sur la machine cible. Cela rend alors l'application plus facile à distribuer et à utiliser pour les utilisateurs qui n'ont pas de connaissances techniques spécifiques. Le logiciel est aussi protégé au niveau de son code source. L'exécutable est compilé, ce

qui signifie qu'il offre des meilleures performances d'exécution par rapport aux scripts Python interprétés. En effet, il est pré-compilé et optimisé pour de l'exécution directe.

Nous avons donc obtenu l'interface ci-dessous, qui correspond ici au menu principal et à interface pour importer un audio à analyser :



Enfin, la page qui affiche la suspicion d'être atteint de TSA, avec le spectrogramme correspondant est telle qu'elle pour un enfant n'étant pas atteint de TSA :



Pour conclure ce dernier chapitre, nous sommes donc parvenus à développer un outil fonctionnel qui pourrait être utilisé par les médecins, en complément des méthodes de diagnostic actuelles. L'ensemble des techniques utilisées pour développer ce projet se sont ainsi articulées autour des caractéristiques vocales exploitables. Tous les thèmes se sont donc emboîtés les uns par rapport aux autres afin de concrétiser notre projet, l'aide au diagnostic précoce des TSA.

## Conclusion technique

En somme, le projet que nous avons pu développer prouve la possibilité de développer des moyens technologiques non invasifs afin d'initier un diagnostic de TSA avant l'apparition du langage. De plus, cette méthode ne représente pas une contrainte pour l'enfant, car la collecte des audios se fait chez un pédiatre, de la même manière qu'une simple visite médicale. Cette intervention ne fait donc aucunement ressentir à l'enfant une pression psychologique liée à une différence de traitement. De même qu'un suivi pédiatrique habituel, l'enfant sera régulièrement enregistré dans différentes situations, afin d'assurer au mieux la fiabilité des suspicions.

Ainsi, la mise en place de Prosod'IA permet d'assurer une prise en charge précurseuse de l'enfant, et donc de minimiser l'impact de ces troubles autant dans sa vie d'enfant que dans sa vie en société. Après avoir échangé avec des acteurs concernés, nous pouvons envisager des perspectives de partenariat avec la Cohorte Marianne pour obtenir les droits pour un meilleur protocole, mais aussi pour permettre d'étendre l'application sur le long terme dans l'utilisation médicale quotidienne.

La réalisation de ce projet a nécessité de nombreuses compétences techniques pour rebondir et pallier chaque problème et s'approcher de l'objectif final. Nous sommes ainsi parvenus à apporter une application d'aide au diagnostic fiable, et à démontrer la faisabilité de cette aide au diagnostic, en atteignant une précision au-delà de 90%.

Par ailleurs, nous avons pris connaissance des précédents travaux effectués sur ce sujet, notamment celui d'un ancien groupe d'ESIEE Paris. Notre projet s'inscrit donc dans sa continuité, ayant obtenu des résultats globalement plus satisfaisants, passant de 86% à 93% de validité. Toutefois, malgré des premiers tests optimistes, notre application possède encore certains points d'amélioration. En particulier, la base de données qui nous a été fournie contient de nombreuses failles.

En premier lieu, il aurait fallu agrandir la base de données pour pouvoir étendre les entraînements de l'IA. Cela permettrait ainsi à l'IA de différencier plus facilement les enfants sains des enfants atteints de TSA, en étudiant un plus large panel de possibilités et de cas distincts.

En second lieu, la précision et la fiabilité des résultats générés par l'application peuvent avoir des conséquences importantes sur le diagnostic et le traitement des TSA. Pour cela, nous devons obtenir une base de données avec un meilleur protocole de diagnostic. Il doit passer par un comité de protection des personnes, ce qui nous permettrait d'être sûrs que les familles aient bien donné leur accord pour utiliser les enregistrements à des fins de recherche.

Bien qu'actuellement nous ayons une base de données de vocalises anonymes, qui ne causerait donc pas de soucis par rapport au consentement, nous devons garder un œil éclairé sur quelques notions de respect des pratiques. Ainsi, une base de données avec un protocole fiable et encadré nous aurait permis de prétendre à une application réellement fonctionnelle.

L'obtention d'une base de données fiable et pertinente est donc la clef de l'amélioration de ce projet.

Si nous avions disposé d'une base de données comprenant un protocole de diagnostic, il aurait également été nécessaire de prendre en compte les problèmes de sécurité des données. En effet, notre base de données est pour le moment composée de vocalises anonymes, il n'est pas pertinent de

prendre en compte la protection des données. Toutefois, en disposant d'une véritable base de données avec un protocole de diagnostic, il aurait été nécessaire de mettre en œuvre des outils pour garantir la sécurité des données. Cette question est un point important, sachant qu'il s'agit de données médicales sensibles et devant donc rester confidentielles.

En outre, cela s'inscrit dans une logique d'amélioration de la précision du résultat. À ce jour, malgré les défauts de la base de données, les résultats restent encourageants pour de futures améliorations. De tels résultats ont été obtenus par l'exploration de nombreux modèles d'intelligence artificielle, mais l'auto-encodeur, que nous n'avons pas pu exploiter, a un énorme potentiel pour une base de données plus propre. De plus, nous pourrions également envisager d'utiliser une intelligence artificielle par tranche d'âge. Un tel procédé pourrait donc augmenter la précision d'aide au diagnostic.

## Bilan du projet

### Partie 1 : Référentiel compétences de l'ESIEE

#### 1.1- Concevoir des systèmes, applications ou solutions d'ingénierie complexes

Au cours du projet, nous avons pu développer la compétence de conception des systèmes et solutions d'ingénieurs complexes. Cette compétence permet de proposer des solutions pertinentes en prenant en compte les problématiques, les besoins, les contraintes et en adoptant une méthodologie appropriée.

Effectivement, dans le cadre du projet, nous avons dû prendre en compte les diagnostics actuels afin de connaître les besoins des médecins et les outils dont ils manquent. En tant qu'ingénieurs, l'objectif était d'apporter une solution innovante à un problème médical encore trop peu étudié, dans le but d'épauler les médecins et de leur proposer un outil supplémentaire pour faciliter un aspect du diagnostic des TSA.

Pour développer cette application innovante, il a fallu s'adapter aux éléments pertinents à considérer. Toutes les caractéristiques vocales ont été étudiées, et selon les contraintes de l'IA ou des moyens disponibles, nous avons gardé uniquement les plus exploitables. Une fois le travail en amont réalisé, nous avons pu développer une méthodologie adaptée au projet.

Ainsi, cette expérience nous a appris à intégrer les contraintes contextuelles et à évaluer la pertinence de différentes solutions. Professionnellement, cela nous a appris à faire des études de faisabilité et à élaborer des applications et des solutions d'ingénierie.

#### 1.2- Réaliser des systèmes, applications ou solutions d'ingénierie complexes

Grâce à ce projet, nous avons également pu développer la compétence de réalisation des systèmes et applications ou solutions d'ingénierie complexes. Elle réside dans la réalisation d'un prototype du système en rendant fonctionnel le système.

Les composantes essentielles font partie de l'optimisation des performances, ici en développant une IA capable de fournir une aide au diagnostic la plus fiable possible. Cette fiabilité permet donc

d'assurer la qualité du système, en validant les solutions mises en œuvre et en intégrant les évolutions technologiques.

Dans ce contexte, nous avons intégré la combinaison de deux IA : MobileNet et VGG-16, et les solutions mises en œuvre ont permis d'approcher une précision de diagnostic de 90%.

Afin de rendre le système fonctionnel, nous avons longuement recherché le meilleur réseau de neurones sur lequel se baser, et c'est à la suite de nombreux tests et entraînements de l'IA que nous avons pu obtenir de si bons résultats. En effet, le travail de recherche effectué sur l'architecture et le modèle a conduit au développement d'une solution, tout en combinant les résultats de ces recherches de manière homogène.

Ainsi, cette expérience nous a appris à choisir ou à adapter un composant d'une application ou d'une solution d'ingénierie complexe. Puis, à tester, à valider et à intégrer ces composants dans des applications ou des solutions d'ingénierie plus complexes. Professionnellement, cela nous a appris à mettre en œuvre des solutions opérationnelles et optimisées répondant aux besoins, tout en intégrant des systèmes existants pour constituer un système plus global.

### 1.3- Gérer un projet

Par ailleurs, nous avons appris à gérer un projet en identifiant et en prenant en compte les risques et les contraintes, en coordonnant efficacement l'équipe, en mettant en valeur notre projet et en suivant une méthode de gestion de projet.

Pour ce faire, nous avons notamment utilisé des outils découverts en début de semestre, tels que le Gantt [cf. annexe 5] et la matrice des risques [cf. annexe 6]. De plus, afin de gérer au mieux la répartition des tâches et l'avancée du projet, nous avons assigné le rôle de Chef de Projet à Lila Razani. Elle nous a permis de coordonner efficacement l'équipe, par le biais d'un Trello par exemple et d'autres outils de communication.

Après avoir évalué les risques, le Gantt nous a permis de gérer notre temps et de se fixer des objectifs hebdomadaires. En outre, chaque semaine, le rendez-vous avec notre tutrice nous a offert un cadre et cela nous a guidés dans la mise en place d'une méthode de gestion de projet.

Pour conclure, cette expérience nous a appris à définir les objectifs, les tâches et les étapes du projet. Nous avons su affecter les ressources aux tâches en fonction des contraintes et nous adapter aux aléas du projet. En effet, des problèmes techniques tels que la qualité de la base de données n'ont pas constitué de réels obstacles. Nous avons su réagir et mettre en place une solution immédiate.

Enfin, les rendez-vous hebdomadaires nous ont appris à nous rendre compte des avancées et du résultat du projet et à contribuer efficacement à la dynamique du groupe.

Professionnellement, le projet nous a ouvert des compétences dans le pilotage d'une équipe pour la réalisation d'un projet, l'identification et la mise en place des ressources matérielles et la présentation et mise en valeur du projet.

## 1.4- Agir avec une démarche scientifique et éthique

Tout au long du projet, nous avons pu agir dans une démarche scientifique et éthique. Nous avons adopté une méthodologie rigoureuse et constructive, en faisant preuve d'autonomie, en actualisant nos connaissances, en assumant la responsabilité de nos décisions et en mesurant les impacts humains et légaux.

Afin d'accomplir cela, nous nous sommes tournés vers une méthodologie en deux phases. La première phase était celle de recherche approfondie. Bien qu'elle nous ait pris du temps, elle nous a permis de passer à la phase technique plus confiants et en ayant une plus ample connaissance du sujet. Ainsi, nous avons pu travailler de manière autonome et prendre des décisions de nous-même étant donné que nous étions bien informés sur tous les aspects du projet.

Nous avons mis en œuvre notre projet en prenant compte les impacts humains et légaux. En outre, nous avons soulevé toutes les questions éthiques essentielles liées à la problématique, en les reportant notamment dans un rapport dédié. Notre projet se concentrerait donc sur des questions de fiabilité de diagnostic et de sécurité des données.

En somme, cette expérience nous a donné l'occasion de formuler une problématique, d'expliciter des préoccupations et des raisonnements liés à l'éthique et à la responsabilité sociétale, d'apporter des arguments pour justifier les choix et d'adopter des méthodes permettant de résoudre un problème. En effet, le sujet sur lequel le projet est centré reste très peu développé, et de surcroît, toutes les problématiques éthiques et sociétales doivent être étudiées de près. Celui-ci touchant au domaine médical, nous nous devons d'être des plus rigoureux.

Sur le plan professionnel, nous serons ainsi en mesure d'interpréter des résultats, d'analyser et faire des évaluations comparatives, tout en prenant des décisions en tenant en considération les avancées technologiques, réglementaires et des considérations éthiques et sociétales. Ces compétences et ces tâches à accomplir, ont été utilisées tout au long de la réalisation du projet. Compétences que nous avons également acquises et renforcées.

## Partie 2 : Bilan personnel

### 2.1- Retour d'expérience de Lila

Au cours de ce projet, j'ai pu développer mes compétences en gestion de projet. En effet, cela m'a permis de comprendre les difficultés à s'adapter aux autres et à leur rythme. Alors que la répartition des tâches s'est faite naturellement, j'ai été assignée à celles à réaliser en premier, ainsi j'ai dû faire preuve d'efficacité et de réactivité pour permettre aux autres membres de continuer sur la suite du projet. De plus, j'ai appris à travailler en groupe étant donné que nous nous entraînions beaucoup, afin de respecter les dates de rendu et les objectifs hebdomadaires. Pour cela, j'ai compris l'importance de la communication et de l'entraide.

Par ailleurs, ce projet m'a également apporté technique et rigueur. Le travail effectué résidait principalement dans de la recherche, et il ne suivait donc pas un cheminement linéaire. Cela m'a permis de travailler autrement, avec patience et persévérance, afin de rebondir sur chaque erreur et de proposer des solutions.

De plus, travailler sur un projet dans le domaine de la santé était mon souhait depuis le début, j'ai donc immédiatement trouvé le sujet très intéressant et cela m'a permis de m'impliquer davantage dans les aspects plus techniques. En outre, j'ai appris à utiliser Matlab et cela m'a permis de mieux comprendre ce logiciel que je ne maîtrisais pas du tout initialement.

En somme, le projet m'a ainsi ouvert à de nouvelles compétences, aussi bien techniques que relationnelles.

## 2.2- Retour d'expérience d'Anaelle

Durant ce projet, j'ai eu l'opportunité de renforcer mes compétences en travaillant en équipe, en m'adaptant aux autres mais aussi en respectant les échéances que nous avions établies. J'ai également pu renforcer mon autonomie et mon sens de l'organisation du fait que nous avons dû nous-mêmes répartir le travail et fixer nos échéances.

D'un point de vue plus technique, j'ai pu développer de nouvelles compétences en programmation. Au début je n'étais pas à l'aise avec la programmation en Matlab et en Python, malgré le fait que nous les avions étudié au cours de l'année. Lors de ce projet, nous avons principalement utilisé Python et Matlab, ce qui m'a permis de mieux me familiariser avec ces deux langages.

Bien que les compétences techniques acquises lors de ce projet ne s'appliquent pas nécessairement à la filière que je souhaite intégrer, Cybersécurité, cela m'a néanmoins apportée des compétences supplémentaires pour mon futur parcours. J'ai également pu constater que même dans un projet relié à la santé, il pouvait y avoir des notions de sécurité, que nous n'avons pas prises en compte lors de la réalisation du projet mais qui peuvent être importantes. En effet, notre projet inclut une base de données et il serait intéressant et important de prendre en compte la protection de ces données.

De manière générale, ce projet m'a donné l'opportunité d'améliorer de nombreuses compétences tout en me rassurant dans mon choix de filière. En effet, l'une des principales raisons qui me pousse à choisir cette filière est que la cybersécurité peut s'appliquer à la plupart des domaines, et j'ai pu en être certaine lors de la réalisation du projet.

## 2.3- Retour d'expérience de Nathan

Travailler sur ce projet dédié au trouble du spectre de l'autisme a été une expérience profondément enrichissante et formatrice. Dès le début, j'étais conscient de l'importance et de la sensibilité de ce sujet. Passionné par l'intelligence artificielle, l'opportunité de développer une application dotée d'IA pour aider les personnes atteintes de ce trouble a été une grande source de motivation.

Cette expérience m'a permis de plonger dans la compréhension des divers aspects du trouble du spectre de l'autisme. J'ai pu approfondir mes connaissances sur les besoins spécifiques des personnes atteintes de TSA et sur les approches les plus efficaces pour les soutenir.

En conclusion, ce projet a été bien plus qu'un simple exercice technique. Il m'a offert une perspective unique sur les défis et les opportunités liés au trouble du spectre de l'autisme. Cette expérience a renforcé ma conviction que l'intelligence artificielle peut jouer un rôle clé dans l'amélioration des conditions de vie des personnes atteintes de TSA. Je suis reconnaissant d'avoir eu l'opportunité de

travailler sur un projet aussi significatif et j'espère continuer à contribuer à des initiatives qui allient technologie et impact social.

#### 2.4- Retour d'expérience de Louka

Au cours de ce projet, j'ai eu l'opportunité de développer mes compétences de recherche et de travail en équipe. Je me suis principalement occupé de la partie Intelligence Artificielle qui est une des dernières étapes du projet et qui nécessite d'avoir une base de données opérationnelle pour fonctionner correctement. C'est pourquoi durant les premières semaines, mon travail portait principalement sur de la recherche afin d'acquérir des connaissances en IA. Ce travail nécessitait beaucoup de rigueur et d'assiduité qui n'étaient pas forcément mes points forts avant ce projet mais que j'ai pu développer.

Par ailleurs, nous étions deux à travailler principalement sur l'IA et nous dépendions en quelques sortes des résultats fournis par nos camarades qui travaillaient sur les étapes précédentes, c'est pourquoi la communication a été très importante afin d'exprimer nos attentes et contraintes pour la base de données ainsi que de leur côté nous dire leur limite. Il a aussi été important de se partager convenablement les tâches afin d'optimiser au maximum le temps passé sur chaque partie du développement de l'IA.

Les compétences techniques acquises durant ce projet s'alignent bien avec mon ambition de poursuivre la filière Data Science et Intelligence Artificielle. J'ai pu approfondir ma compréhension des algorithmes et des modèles de machine learning, en particulier en utilisant des outils comme Matlab et Python pour la mise en œuvre de modèles de classification.

Ce projet m'a donc non seulement permis de développer des compétences techniques spécifiques qui me seront bénéfiques dans mon choix de cursus, mais aussi des compétences primordiales de travail en équipes qui sont très importantes pour le métier d'ingénieur.

#### 2.5- Retour d'expérience de Livio

Dès le début, l'exploration des TSA et la manière dont la technologie peut aider à leur diagnostic précoce m'ont profondément captivé. Cet intérêt était notamment présent grâce à ma prise de conscience que le fait que les TSA n'étaient pas assez connus du grand public. L'objectif de notre projet était de développer une application capable de détecter des indices précoce de TSA à partir des enregistrements vocaux de bébés. Cette tâche impliquait une combinaison de traitement du signal et d'intelligence artificielle, deux domaines dans lesquels je voulais approfondir mes connaissances.

Mon rôle principal dans l'équipe était de me concentrer sur l'extraction de caractéristiques vocales à partir des enregistrements audio. Cela comprenait l'utilisation de techniques comme les MFCC et les spectrogrammes pour capturer les nuances et les caractéristiques uniques des voix des bébés.

Ces techniques ne m'étaient pas familières, mais j'ai su les comprendre et les intégrer dans le panel d'outils à ma disposition. De plus, ce projet a mobilisé des connaissances sur plusieurs langages de programmations, de matlab à python en passant par du C.

Vers la fin du projet, j'ai pu étendre ma contribution en participant à l'entraînement et à la validation des modèles d'intelligence artificielle. Utilisant TensorFlow et Keras, nous avons développé des réseaux de neurones capables de distinguer les enregistrements vocaux de bébés avec TSA de ceux sans TSA. Mon implication a consisté à ajuster les hyper paramètres des modèles et à analyser les résultats pour améliorer la précision et la fiabilité de nos prédictions.

Travailler au sein de notre équipe a été particulièrement enrichissant. Nous avons partagé nos connaissances, résolu des problèmes ensemble et, de nombreuses fois, ajusté nos approches au fur et à mesure que nous progressions dans le projet. Cette dynamique de travail en équipe a non seulement renforcé mes compétences techniques, mais m'a également permis de me conforter dans ma vision du travail d'équipe, notamment en prônant la collaboration et la communication. Les échanges constants avec mes coéquipiers et les spécialistes du domaine ont été cruciaux pour surmonter les défis techniques et conceptuels.

En résumé, ce projet m'a offert une perspective nouvelle sur l'application de la technologie pour résoudre des défis médicaux. Il a enrichi ma compréhension des TSA et m'a donné l'opportunité de contribuer activement à un projet ayant un impact potentiellement significatif sur la vie des enfants et de leurs familles. Cette expérience m'a permis de développer des compétences précieuses en extraction de caractéristiques vocales, en IA et en travail d'équipe, tout en me sensibilisant davantage aux enjeux de la santé infantile.

## 2.6- Retour d'expérience de Tom

Travailler sur le projet d'aide au diagnostic des Troubles du Spectre Autistique (TSA) chez les bébés de moins de 2 ans a été une expérience extrêmement enrichissante. Pendant deux mois, j'ai eu l'occasion de me plonger dans un domaine entièrement nouveau pour moi et de contribuer activement au développement d'un logiciel d'intelligence artificielle dédié à cette cause.

Dès le début, l'exploration des TSA et la manière dont la technologie peut aider à leur diagnostic précoce m'ont profondément captivé. Apprendre et appliquer Python pour créer des modèles d'IA capables d'analyser des données cliniques et comportementales a été à la fois stimulant et gratifiant. C'était une opportunité unique de combiner mes compétences en programmation avec un objectif concret et significatif dans le domaine de la santé.

Travailler au sein d'une équipe collaborative a également été une expérience enrichissante. Nous avons partagé nos connaissances, résolu des problèmes ensemble et ajusté nos approches au fur et à mesure que nous progressions dans le projet. Cette dynamique de travail en équipe a non seulement renforcé mes compétences techniques, mais m'a également permis de découvrir l'importance de la collaboration et de la communication.

Mon attrait pour la création s'est fait ressentir également dans ce projet pour toute la partie graphique du logiciel.

En résumé, ce projet m'a offert une perspective nouvelle sur l'application de la technologie pour résoudre des défis médicaux complexes. Il a enrichi ma compréhension des TSA et m'a donné l'opportunité de contribuer activement à un projet ayant un impact potentiellement significatif sur la vie des enfants et de leurs familles.

## Bibliographie

Autistes sans Frontières. (n.d.). **Site Web.**

Consulté le 27 mai 2024, à partir de

<https://www.autistessansfrontieres.com/>

Femmes autistes francophones. (n.d.). **Contact.**

Consulté le 27 mai 2024, à partir de

<https://femmesautistesfrancophones.com/>

Thales. (n.d.). **HIPAA-HITECH Compliance.**

Consulté le 27 mai 2024, à partir de

<https://cpl.thalesgroup.com/compliance/hipaa-hitech-compliance>

Data and More. (n.d.). **Regulations.**

Consulté le 27 mai 2024, à partir de

<https://dataandmore.com/>

Kiteworks. (n.d.). **Quelles sont les normes de conformité des données ?**

Consulté le 27 mai 2024, à partir de

<https://www.kiteworks.com/fr/plateforme/conformite/conformite-reglementaire/>

Les Silences des Justes. (n.d.). **L'importance d'un diagnostic précoce de l'autisme.** Consulté le 27

mai 2024, à partir de

<https://lesilencedesjustes.fr/>

Vallet, S., & Billard, C. (2009). Autisme et troubles du spectre autistique: Une approche interactionniste. **Enfance**, 2(3), 259-284.

Consulté le 27 mai 2024, à partir de

<https://www.cairn.info/revue-enfance.htm>

Haute Autorité de Santé (HAS). (2018). **Trouble du spectre autistique de l'enfant et de l'adolescent: Recommandations.**

Consulté le 27 mai 2024, à partir de

[https://www.has-sante.fr/jcms/p\\_3113170/fr/autisme](https://www.has-sante.fr/jcms/p_3113170/fr/autisme)

[10]« Pretrained Deep Neural Networks ». Mathwork. Consulté le 10 juin 2024, [En ligne] à partir de

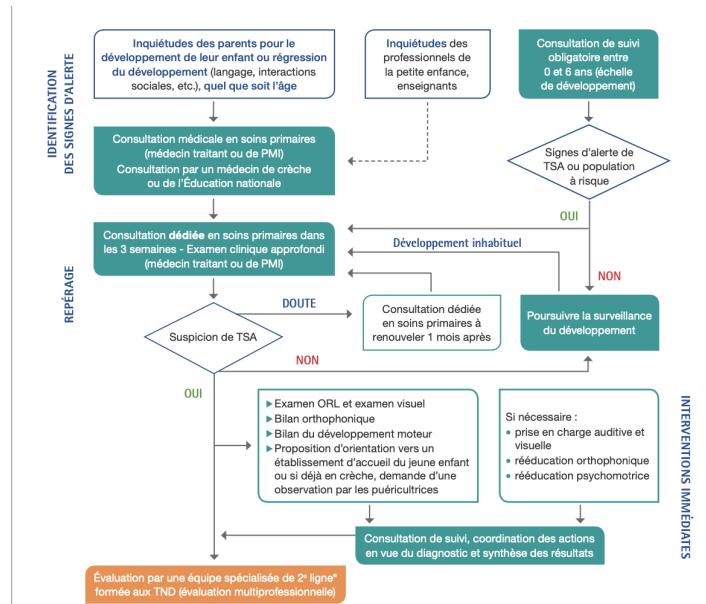
[https://fr.mathworks.com/help/deeplearning/ug/pretrained-convolutional-neural-networks.html?s\\_tid=srchttitle](https://fr.mathworks.com/help/deeplearning/ug/pretrained-convolutional-neural-networks.html?s_tid=srchttitle)

« Transfer Learning : Qu'est-ce que c'est ? » DataScientest. Consulté le 12 juin 2024, [En ligne] à partir de

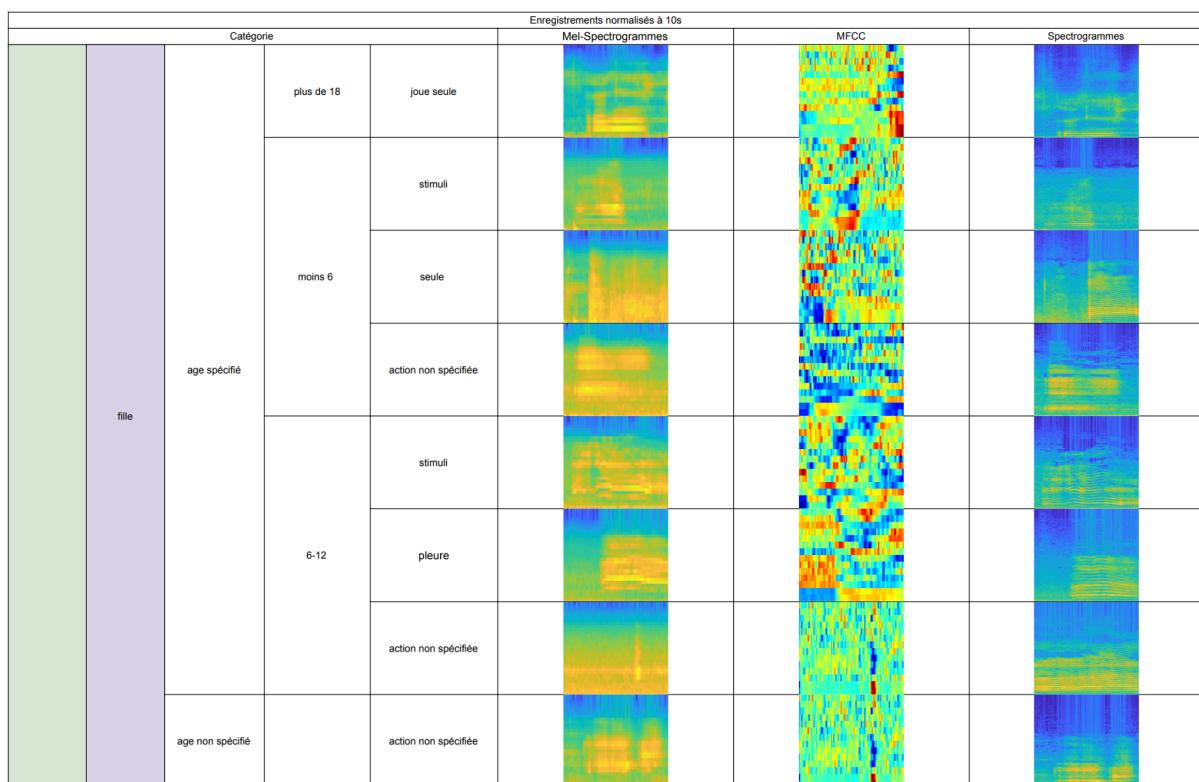
<https://datascientest.com/transfer-learning>

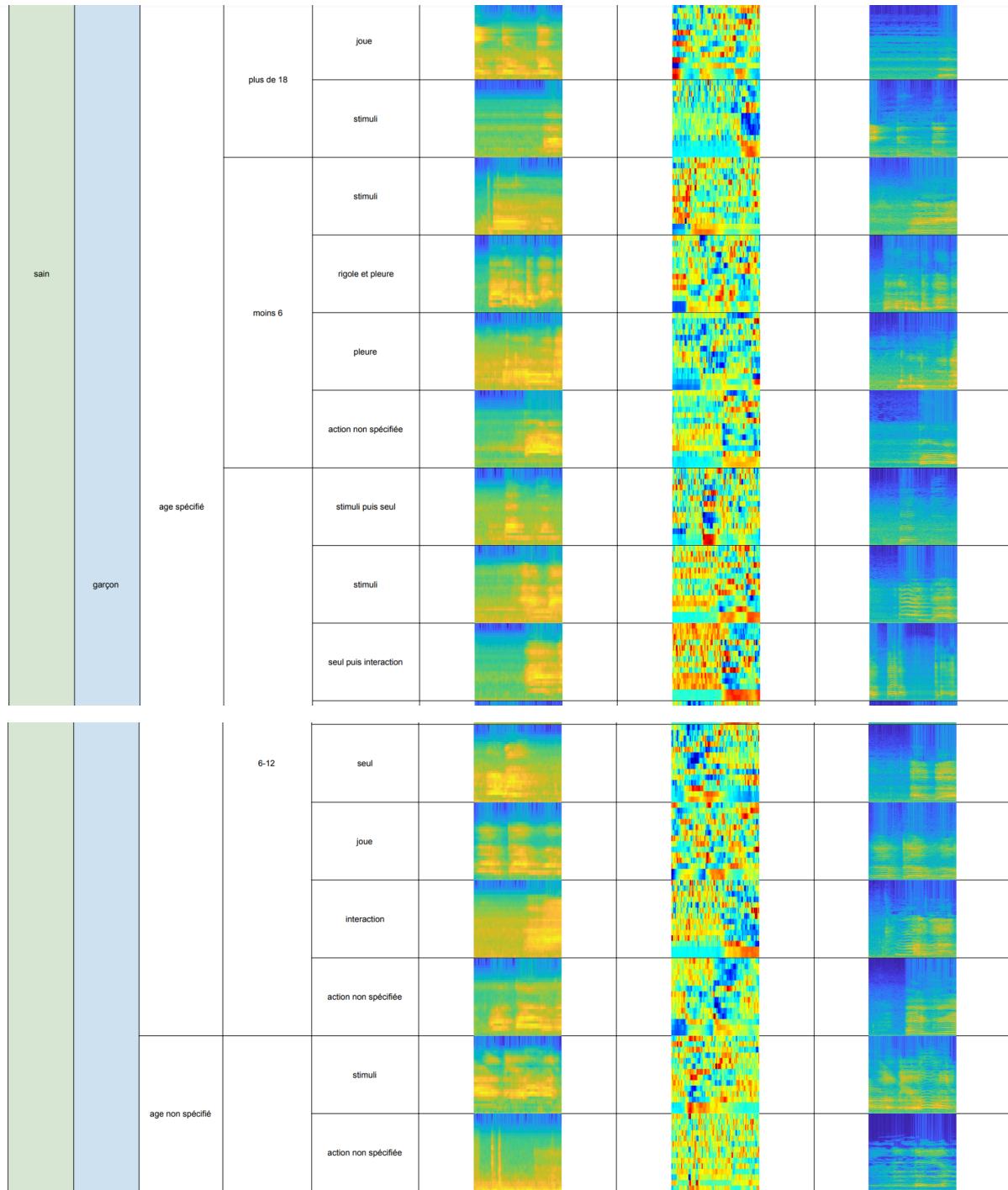
## Annexe

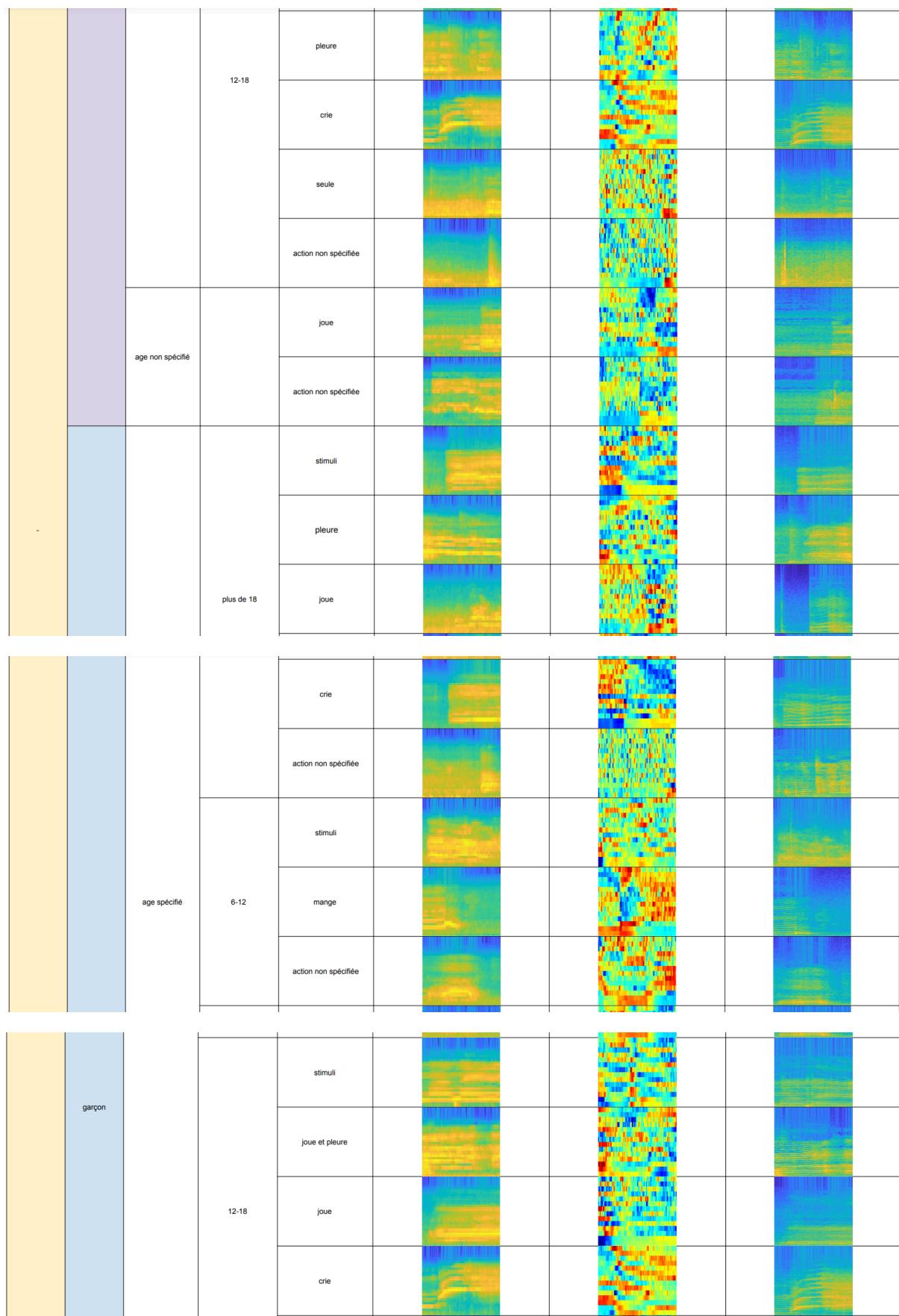
### 1- Schéma du processus de diagnostic actuel :

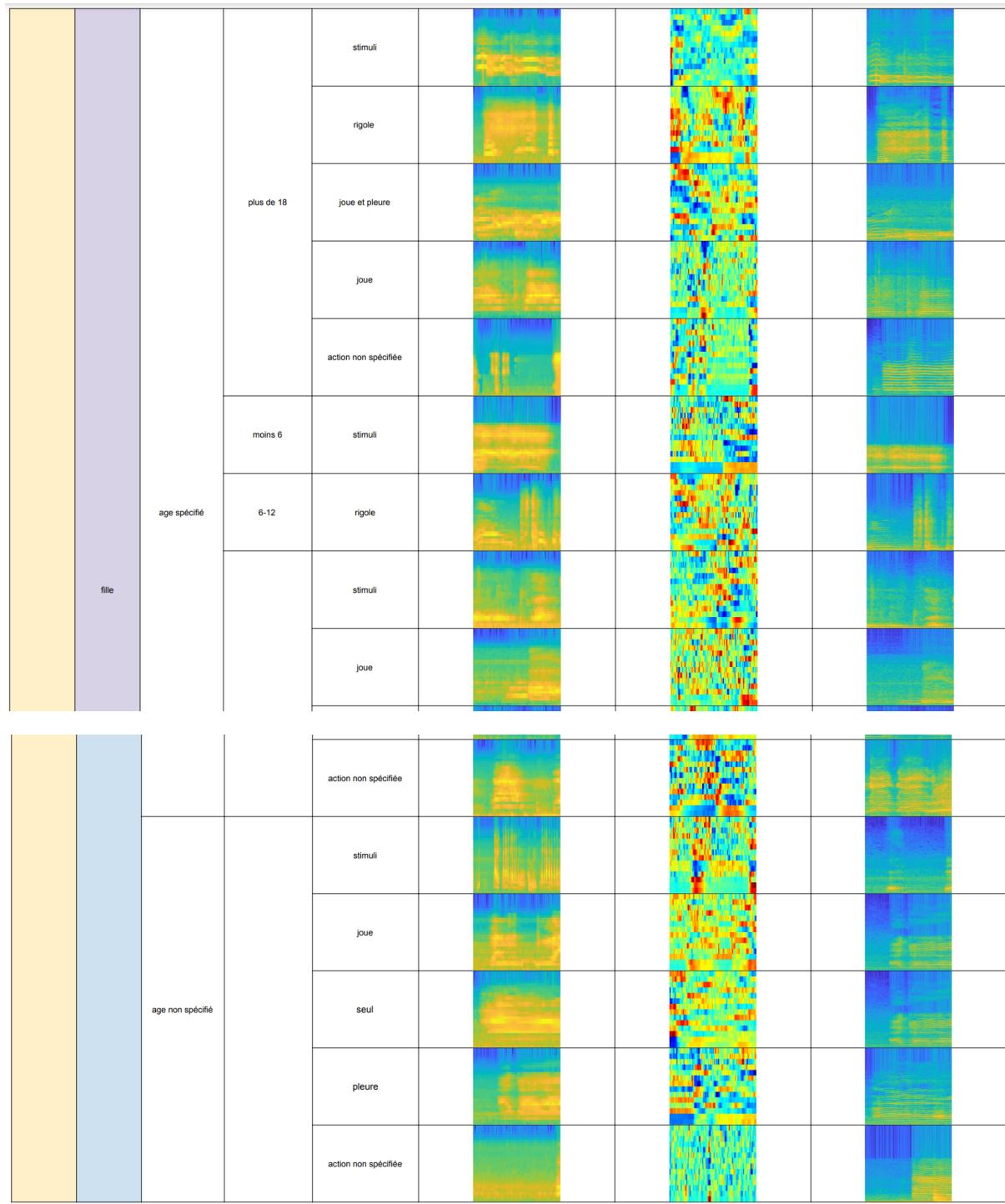


### 2- Tableau de l'étude comparative des images selon chaque condition :









### **3- Extrait de la retranscription de l'interview avec Sandrine Sonié, coordinatrice du centre de ressources autisme pour Rhône Alpes**

#### **Sandrine Sonié**

**“ - Je crois que même les applications actuellement doivent être soumises à un minimum de protocole éthique et puis il faut s'assurer que les bébés sont effectivement autistes, selon quelles méthodes pour le diagnostic.**

*Il faut que vous ayez ces éléments d'abord parce que si finalement c'est pas prêt vous avez pas la certitude du diagnostic derrière on pourrait pas voir si votre appli, votre reconnaissance vocale est bonne.*

*Donc le problème c'est que vous allez passer beaucoup de temps à essayer de faire des filtres et en fait à la base vos données sont pas sûres, pas fiables, enfin elles le sont peut-être mais il faut vraiment vous en assurer avant de vous lancer là-dedans parce que vous avez passé beaucoup de temps à faire du filtrage de de signaux et puis en fait peut-être que les bébés sont autistes mais ils n'ont pas été diagnostiqués selon les recommandations qui sont nécessaires pour la recherche en clinique. Donc vous ne pourrez pas vous en servir, enfin uniquement pour votre école ou votre projet mais ça tombera à l'eau donc faut vraiment vous assurer de la qualité de vos données avant de vous lancer.*

**- C'était surtout pour savoir ce que tu penses du projet et si la prosodie paraît être un indicateur efficace ?**

**- La prosodie toute seule non certainement pas ils ont déjà essayé de faire de l'ittracking tous les jours t'as des trucs diagnostic précoce de l'autisme ça peut être un des facteurs parmi d'autres mais un facteur de prosodie trait traducteur de l'autisme non je pense pas, ça peut être intéressant mais y a beaucoup de bébés en plus qui vont déclencher leur autiste plutôt après 18 mois donc il y en a qui n'ont pas de problème de prosodie y en a qui ont des gros problèmes de prosodie mais c'est pas spécifique à l'autisme hein tous les retards de langage tous les bébés qui ont des problèmes d'audition ils vont avoir des problème de prosodie les gens qui ont d'autres troubles du neurodéveloppement des troubles du développement intellectuel donc la prosodie en elle-même toute seule ne sera pas assez à mon avis prédictive mais ça peut être un des éléments intéressants.**

*Vous avez pas intérêt à dire ça va être une appli pour dépistage de l'autisme par contre vous pouvez dire que vous voulez développer un outil pour retrouver des marqueurs précoces de facteurs de risque tu vois c'est beaucoup plus sobre c'est un marqueur précoce d'un risque d'évolution vers un TSA ou alors même beaucoup parler de troubles du neurodéveloppement comme ça pour vous couvrez et vous avez les TDI des troubles du langage tout ça, là vous serez moins spécifique et donc plus large et surtout vous êtes plus crédibles, parce que de dire que vous allez faire un une appli pour dépistage précoce de l'autisme ça me paraît un peu ambitieux surtout si vous n'avez pas de données fiables derrière.*

*Ce qui est très important c'est de voir dans quelles circonstances ont été recueillies les extraits et quelle est la fiabilité du diagnostic qui a été donné, est-ce que c'est vraiment une équipe pluridisciplinaire et est-ce que ils ont fait partie d'un protocole parce que sinon c'est pas très claire de faire du rétro enfin c'est pas est-ce que ces enregistrements quand ils ont été pris c'était dans le but de retrouver ces éléments-là sinon ça risque de biaiser c'est à dire que par exemple imagine t'as pris*

*des enregistrements pour tout à fait autre chose tu te dis « j'ai envie de voir si y'en a qui sont devenus autistes et je vais voir comment sont leur enregistrement mais finalement ces enregistrements Oh celui-là il me paraît pas très clair celui-ci il est comme ça » et **tu commences à biaiser ton étude parce que l'air de rien tu fais un tri.** Par contre si tu as d'emblée posé ta question dans un protocole de recherche je vais prendre X bébé 100 au départ je vais les suivre je regarde sans préjugé de qui va être autiste ou pas là c'est rigoureux, tu risques pas de biaiser. Si tu as pris des enregistrements pour autre chose et puis finalement tu as décidé de les utiliser quand ça t'intéresse.*

*Peut-être que vous vous devez intervenir juste sur la question du traitement du signal.*

*En tout cas il faudrait que vous récupériez le protocole complet et juste pour voir dans quoi vous vous inscrivez et sinon en gros vous allez passer du temps bah vous sur le plan technique ce serait intéressant mais par contre sur l'objectif final ça peut être vraiment planté si c'est biaisé.*

*- Ce protocole on le retrouve comment ?*

*- C'est votre tutrice qui doit vous le donner.*

*Puis quand même ce serait bien d'être sûr que le protocole soit bien passé à un comité de protection des personnes tu vois c'est un truc de base vous ça vous regarde pas forcément en direct mais c'est toujours bien avant de se lancer sur vos traitements de données être sûr que les familles avaient bien donné leur accord pour qu'on prenne leurs enregistrements je suis sûre que tout ça a été fait mais c'est bien pour vous que vous sachiez que ça existe et que vous posiez ce type de questions.*

*Sur les diagnostics si tu veux je pourrais regarder ce qu'ils ont fait et te dire si oui c'est la bonne méthodologie pour confirmer le diagnostic !*

*- Oui ce serait super merci ! Je ne sais pas comment elle les a retrouvés [les enregistrements] mais si ça avait été repris par internet ou quelque chose comme ça, sachant que c'est déjà en ligne est-ce que la question de consentement est-elle respectée quand même ou pas du tout ?*

*- Alors sur des données anonymes je pense que ça va, par contre après ça me paraît euh après c'est pour prouver le diagnostic si c'est juste des déclaratifs ça suffira pas c'est à dire c'est juste des gens par exemple si des gens on fait sur internet ils ont fait leurs petits enregistrements ils ont dit bah voilà mon enfant est autiste je vous donne ces enregistrements bah on sait pas s'ils sont vraiment autistes.*

*Après si c'est juste une base de données de vocalise je pense que là vous avez pas de souci par rapport à ça mais c'est quand même toujours bien de savoir au moins ce qui a été fait le type d'informations qui a été donné aux familles et simplement des fois ils ont pas besoin de consentement pour les familles mais ils ont au moins déclaré l'étude quelque part ou à moins que ce soit fait en soins courants enfin c'est pas mal d'avoir quand même quelques notions de respect des pratiques parce que tu vois ça c'est souvent en plus c'est des labos de neurosciences ils peuvent être un peu négligents là maintenant c'est moins le cas, on leur demande moins de choses qu'avant donc je pense que maintenant ils sont dans les clous mais quand même pour vous et pour votre projet c'est bien d'être sûr.*

*Demande est ce que en fait normalement, il y a pas forcément besoin d'un avis du CPP mais au moins elle peut vous dire si cette étude relevait de la loi Jardé tu vois t'as une loi de protection des*

**personnes voilà c'est juste vous savoir à quel registre appartient cette étude est-ce que ça relève d'un CPP ou pas juste d'avoir au moins la réponse ça c'est un point de détail.**

- *On a déjà effleuré la question, mais est-ce que tu penses que ça pourrait poser des problèmes d'éthique ?*

- *Ça peut poser en fait sur le repérage précoce ça pose pas de problème sauf si votre outil est biaisé là ça peut poser des problèmes.*

*Là vous allez utiliser des données rétrospectives donc même si l'étude a été faite en prospective y a pas de conséquences pour les personnes sur votre phase de mise en place de votre logiciel par contre si votre logiciel est mal calibré vous dépistez, vous informez trop des familles par défaut tu vois ça peut poser des ça peut causer du stress donc là les risques c'est sûr comme pour tous les outils de dépistage c'est que ça peut stresser des familles.*

- *Finalement quelle perspective d'amélioration tu verrais pour l'application hormis le fait que le critère ne serait pas suffisant dans l'application elle-même ?*

- *Il faudrait que votre application puisse s'intégrer à un ensemble plus vaste d'outils d'aide au diagnostic et donc il faudra effectivement faire avec l'IA on peut croiser des données d'anamnèse avec les facteurs de risque avec d'autres éléments de facteurs de risque et ça peut être un élément en plus.*

*Ce qui peut être amélioré c'est vraiment que cette application soit en quelque sorte toutes les données issues de cette appli soient interopérables ou vous puissiez mettre des questionnaires en parallèle ça peut être intéressant des questionnaires qui sont assez basiques. Nous on a une étude qui s'appelle kit Kat où on dépiste les enfants à 18 mois/24 mois avec un questionnaire qui est pas mal et des questionnaires qui sont déjà validés vous pourriez dire que vous vous rajoutez ces questionnaires à votre appli par contre il y a peut-être des questions de droit d'auteur.*

- *Est-ce que c'est possible de les récupérer sans demander l'autorisation ?*

- *Il faudra faire des droits d'auteur mais vous pouvez pour votre projet vous pouvez dire qu'il faudra peut-être les joindre pour croiser vos données avec des questionnaires.*

*Vous avez toujours intérêt à dire que votre facteur de risque sorti d'un contexte développemental et d'éléments cliniques n'a pas de valeur. C'est comme une donnée d'imagerie quelqu'un qui n'a pas de signe clinique et qu'on trouve un truc sur l'imagerie ça n'a pas de valeur si t'as pas la clinique avec.*

*Donc c'est pour ça qu'il faudra toujours bien faire attention qu'il y ait dans votre appli des données de développement avec les facteurs de risque, les fameux drapeaux rouges de risque ça il faudrait les joindre à votre appli. Car un bébé qui va très bien et qui a juste une prosodie ça va peut-être être un risque faible quoi par contre un bébé qui a eu plein de facteurs de risques à la naissance qui a déjà un frère autiste et cetera plus une prosodie là il y a peut-être plus à ça peut avoir un autre un autre poids c'est pour ça qu'il faut croiser vos données votre donnée isolée ne servira à rien.*

*C'est un facteur de pondération dans un contexte de facteurs de risque et si vous avez un bon outil je dirais. Si ça marche bien vous pourriez proposer votre outil à la cohorte Marianne, notre cohorte, où on suit des bébés dès la grossesse et ça va être des bébés à risque d'autisme et vous pourriez, si vous avancez rejoindre la cohorte elle va durer 10 ans.*

*Elle dure 4 ans en repérage mais ça vous permettrait d'enregistrer les bébés de la cohorte.*

*Souvent on retrouve très peu de choses avant 12 mois quand même dans le développement des différences. Actuellement il y a peu de d'éléments avant 12 mois.*

*- Même dans la prosodie même tout ça c'est toujours après 12 mois en ce moment ?*

*- Souvent dans les études en tout cas au niveau développemental on voit pas grand-chose de différent après on voit les courbes entre les 2 populations ceux qui vont devenir autistes et pas autistes qui se séparent à partir de 12 mois c'est là que ça commence à être plus clair.*

*- OK d'accord bon bah merci beaucoup en tout cas pour tout ça il y a beaucoup de d'infos il y a pas mal de points c'est pas évident*

*- Mais bon vous vous intéressez aux aspects techniques mais c'est bien d'avoir quand même la vision globale de dans quoi vous vous inscrivez.*

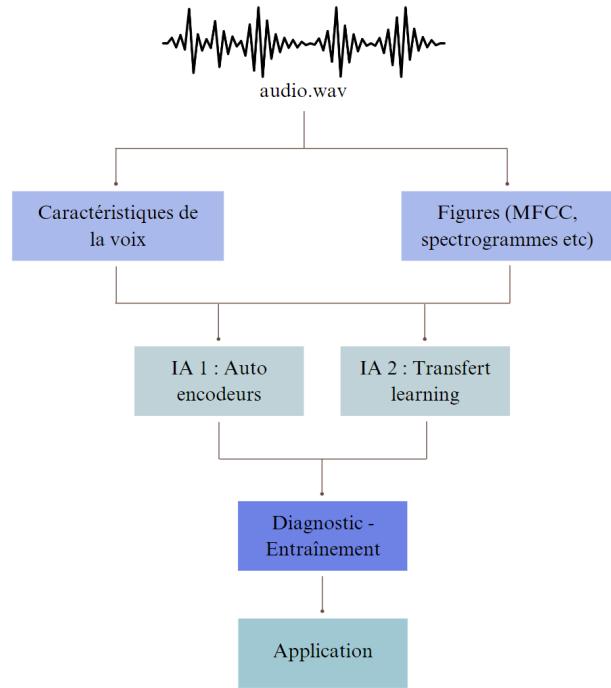
*- Oui c'est sûr bah c'est pour ça que qu'on a fait appel à toi parce qu'on s'y connaît pas trop.*

*- Récupère déjà ce que tu as et puis tu peux revenir vers moi et quand comme ça je pourrais un peu mieux te guider sur ce qui ce qui paraît pertinent ou pas.*

*- Ok merci beaucoup je te dirai où est-ce qu'on en est je reviendrai vers toi.*

*- Super merci à bientôt au revoir ! “*

#### 4- Organigramme du projet



#### 5- GANTT

Progrès/Tâche	Mai				Juin			
	1	2	3	4	5	6	7	8
Définir le problème	■							
Bibliographie	■	■■■						
Isoler les voix des bébés			■■■■					
Scalogramme et spectrogramme			■					
Pitch			■■					
MFCC et formants			■■■■					
Utilisation RN pré-entraîné				■■■■■■				
Autoencoder				■■■■■■				
Phase d'entraînement et tests					■■■■■■■■			
Développement application						■■■■■■■■		

## 6- Matrice des risques

