Stage de master 2027

Développement d’un outil logiciel d'aide à l'interprétation de données médicales de Doppler transcrânien et des différentes représentations issues de modèles d'apprentissage

Résumé - Développement Python, visualisation interactive de données médicales et intelligence artificielle
Ce stage, mené au sein du laboratoire CREATIS (INSA Lyon) en collaboration avec l’entreprise Atys Medical, porte sur le **développement d’un outil interactif de visualisation et d’analyse de données médicales**. L’objectif est de faciliter l’exploration de signaux Doppler transcrâniens (TCD), utilisés pour détecter les emboles cérébraux, et d’analyser les performances de modèles de classification par apprentissage profond (deep learning).

Le ou la stagiaire développera un tableau de bord interactif en Python permettant de visualiser les données sous forme de nuages de points, d’explorer les spectrogrammes et les métadonnées, de filtrer les données et de comparer les résultats des modèles. L’outil sera notamment **testé sur des données cliniques** provenant d’une collaboration internationale avec l’Institut de Cardiologie de Montréal, afin d’étudier les différences entre centres hospitaliers et protocoles d’acquisition.

Ce stage offre l’occasion de travailler sur des problématiques concrètes de data science, visualisation interactive, analyse exploratoire de données et intelligence artificielle appliquée à la santé, dans un environnement associant recherche académique, développement logiciel et collaboration industrielle.

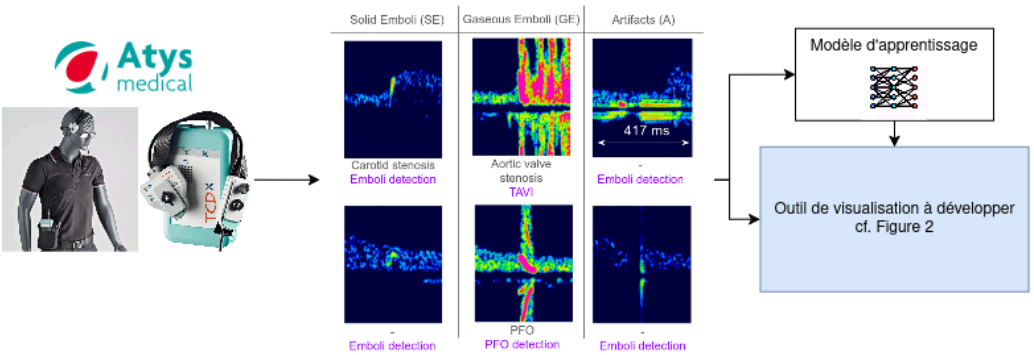


Figure 1. Résumé global de la pipeline avec des exemples de spectrogrammes issus du TCD-X d’Atys Medical

Contexte
Les emboles cérébraux sont des particules gazeuses ou solides circulant dans les vaisseaux du cerveau. Ils peuvent notamment être à l’origine d’accidents vasculaires cérébraux (AVC), qui comptent parmi les principales causes de décès et de handicap dans le monde [1]. Le Doppler transcrânien (TCD) est la seule technique non invasive permettant

de détecter et de surveiller en temps réel les emboles circulant dans les artères cérébrales, sur des durées prolongées. La figure 1 présente quelques exemples de spectrogrammes de signaux TCD.

CREATIS est un laboratoire de recherche en imagerie médicale situé à Lyon et **Atys Medical** est une entreprise située dans la région lyonnaise qui conçoit et commercialise des dispositifs médicaux, notamment un TCD portable et robotisé.

Durant le projet qui nous réunit, des modèles de classification par apprentissage automatique profond [2] ont été développés et partiellement intégrés au logiciel de l'entreprise. Nous nous intéressons maintenant à **l'amélioration de ces modèles dans le temps**, à la fois lors d'échanges de données avec des hôpitaux pour des études cliniques, mais aussi dans lors de leur utilisation en routine dans les hôpitaux où ils sont déployés. Nous avons notamment récemment rejoint une étude avec **l'Institut de Cardiologie de Montréal**, qui nous donne accès à des données TCD ainsi qu'aux métadonnées associées. Ces données constituent une nouvelle source d'information pour étudier la diversité des signaux et des contextes cliniques rencontrés.

L'objectif de ce stage est de développer un **outil de visualisation et analyse des données** pour l'aide à l'interprétation des données médicales, et plus particulièrement pour comparer les données entre centres et analyser qualitativement les résultats des modèles développés. L'élément principal est la visualisation des données dans un espace de représentation commun via un nuage de point, qui puisse permettre d'observer directement comment les données issues d'un nouveau centre/protocole/service se projettent dans cet espace par rapport à la base de données initiale dont nous disposons. Cette visualisation permet d'enrichir les discussions avec les équipes cliniques, aussi bien au sein de CREATIS qu'avec Atys Medical. Le logiciel sera donc valorisé par l'entreprise et pour l'étude clinique en cours avec le Canada.

Le développement devra s'intégrer à l'environnement logiciel existant, c'est-à-dire une extension du logiciel de l'entreprise fondée sur **Python**. Le choix de la technologie de visualisation reste ouvert ; Plotly+Dash constituent notamment des pistes possibles.

Objectifs et fonctionnalités :
L'outil développé devra notamment permettre :

- **La visualisation interactive des données** sous la forme d'un nuage de points, avec la possibilité d'afficher les images ou spectrogrammes associés à un point ou à un ensemble de points, afin d'explorer les données et leur voisinage.
- **L'exploration des informations associées aux données**, notamment les métadonnées disponibles, les prédictions des modèles et leurs incertitudes.
- **Le filtrage interactif** des données affichées selon différents critères.
- **La visualisation d'histogrammes et de statistiques descriptives**, notamment à l'échelle des sujets.
- **L'extension éventuelle à la comparaison de modèles**, afin d'étudier les différences de prédictions et de comportements sur un même jeu de données.

L’outil devra être conçu de manière modulaire afin de faciliter son utilisation sur de nouveaux jeux de données et son évolution ultérieure.

La figure 2 est un schéma de la pipeline globale.

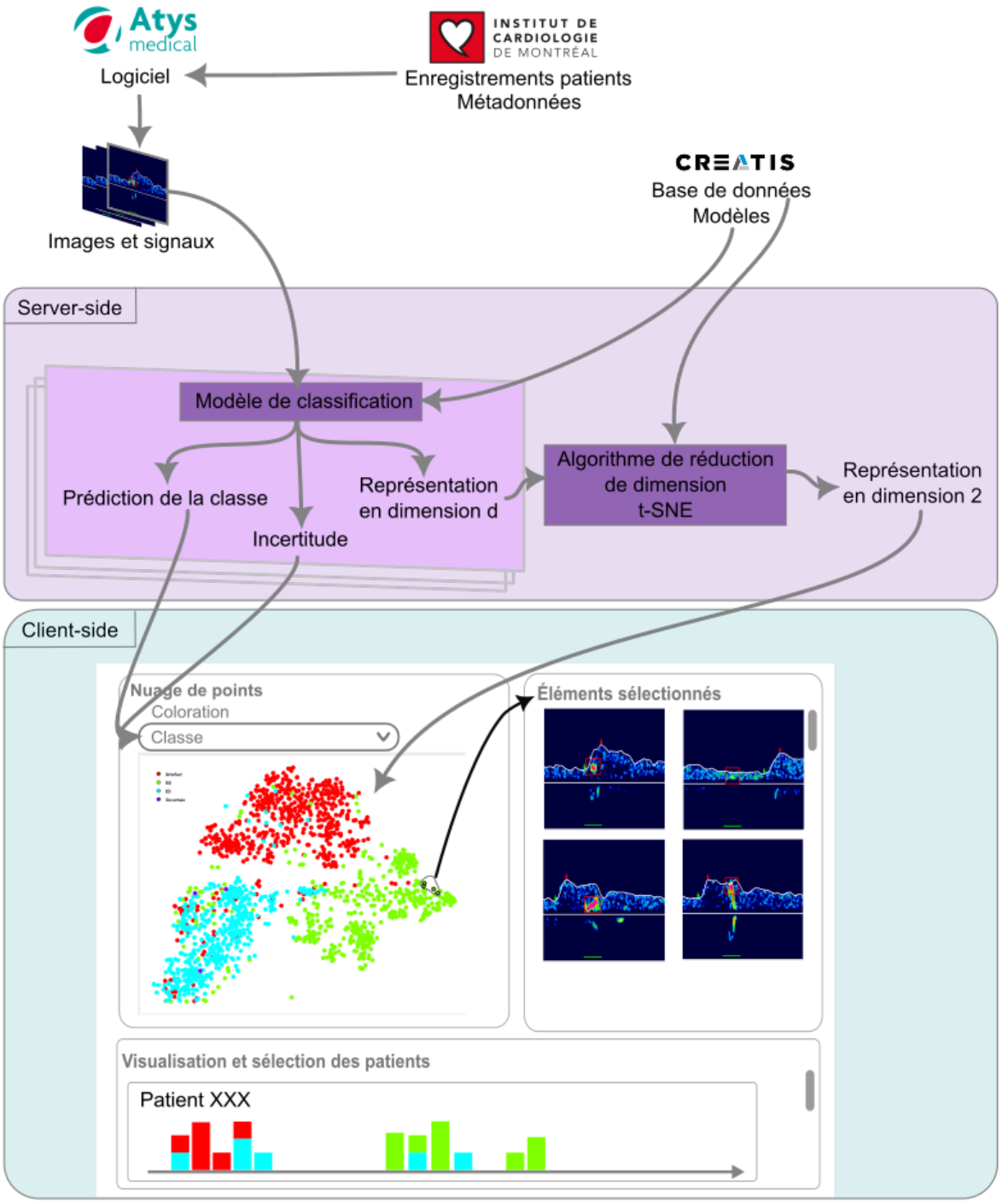


Figure 2. Schéma des informations données en entrée par les différents acteurs, leur traitement, les sorties du traitement et leur visualisation dans l’interface.

Déroulement du stage :
Les principales étapes proposées sont les suivantes :

1. Préciser le cahier des charges et les besoins avec CREATIS, Atys Medical et, si nécessaire, l'équipe de l'Institut de Cardiologie de Montréal.
2. Choisir la technologie de visualisation et définir les modalités d'intégration dans l'environnement logiciel existant.
3. Extraire, structurer et harmoniser les données canadiennes ainsi que leurs métadonnées.
4. Développer le tableau de bord interactif : nuage de points, visualisation des signaux, exploration des métadonnées, filtres et histogrammes.
5. Tester l'outil sur les données canadiennes et réaliser une analyse statistique exploratoire des différences observées entre les jeux de données.
6. Généraliser la chaîne de traitement afin de faciliter l'intégration de nouveaux jeux de données et assurer la modularité de l'ensemble.
7. Documenter le code et les modalités d'utilisation de l'outil.

Ces étapes pourront être adaptées en fonction des besoins identifiés, de l'avancement du projet et des propositions du ou de la stagiaire.

Les délivrables attendus :

- Le code source documenté de la chaîne de traitement et de l'outil de visualisation, accompagné des instructions nécessaires à leur installation et à leur utilisation.
- Une version exécutable ou facilement déployable de l'outil, selon les contraintes de l'environnement logiciel cible.
- Un rapport présentant l'analyse exploratoire et statistique des données canadiennes, les principales observations et les limites de l'étude.

Profil recherché

Étudiant en master avec un intérêt pour le développement logiciel, la visualisation interactive, le traitement et la gestion de données.

- Compétences techniques : Python et Git/GitHub. Une expérience en développement d'interfaces utilisateur ou en visualisation de données est appréciée.
- Connaissances scientifiques : une compréhension de l'apprentissage automatique est un atout.
- Langues : maîtrise de l'anglais nécessaire pour lire et rédiger de la documentation technique et scientifique, ainsi que pour participer aux présentations et échanges au laboratoire.
- Motivation : intérêt pour les applications médicales, le travail interdisciplinaire.

Remarque : le ou la stagiaire effectuera son stage au sein du laboratoire CREATIS rattaché à l'INSA Lyon, avec potentiellement une ou deux visites à l'entreprise et des points réguliers. L'environnement sera donc celui d'un laboratoire de recherche académique.

Supervision

Mathilde Dupouy (doctorante)
Philippe Delachartre (Professeur des Universités)
Blaise Kévin Guépié (Maître de conférences)
Yamil Vindas Yassine (Post-doc)

Informations pratiques

Site : Laboratoire CREATIS, campus de la DOUA, Villeurbanne (Lyon)

Gratification : ~650 €/mois

Duration : ~6 mois, début à partir de janvier ou plus (3 semaines de fermeture en août)

Langage : Français ou anglais

Contact : envoyer CV, lettre de motivation et dernier relevé de notes à mathilde.dupouy@creatis.insa-lyon.fr

Bibliographie

[1] J. D. Steinmetz *et al.*, “Global, regional, and national burden of disorders affecting the nervous system, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021,” *Lancet Neurol.*, vol. 23, no. 4, pp. 344–381, Apr. 2024, doi: 10.1016/S1474-4422(24)00038-3.

[2] Y. Vindas, E. Roux, B. K. Guépié, M. Almar, and P. Delachartre, “Guided deep embedded clustering regularization for multifeature medical signal classification,” *Pattern Recognit.*, vol. 143, p. 109812, Nov. 2023, doi: 10.1016/j.patcog.2023.109812.

[3] M. Dupouy *et al.*, “Integrating Uncertainty and Latent Feature Diversity for Robust Memory Replay in Continual Medical Imaging Learning,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2026 Workshops and Challenges*, Springer Nature Switzerland, 2026.