

## Aide au diagnostic : comment prendre en compte les variables de nuisance ?

**Equipe et laboratoire d'accueil** : équipe IMAGEs <http://images.icube.unistra.fr/> au sein du laboratoire ICube (UMR 7357), site d'Illkirch, université de Strasbourg, CNRS.

**Période** : 6 mois à partir de janvier 2019

### Compétences souhaitées

- mathématiques appliquées (statistique, analyse de données, apprentissage)
- programmation (Python de préférence).

### Sujet :

De nombreuses bases de données d'images médicales sont désormais disponibles et offrent la possibilité de développer des méthodes de classification (ex : aide au diagnostic) ou de régression (ex : prédiction du taux de survie) à partir d'un nombre très important d'échantillons. Cependant, ces données sont généralement corrompues par des variables de nuisance telles que l'âge, le sexe et les paramètres d'acquisition qui peuvent détériorer les performances des méthodes d'apprentissage. L'objectif du stage est de développer des méthodes qui permettent de supprimer les effets de ces variables. Cela revient à estimer pour chaque sujet les attributs qu'il devrait avoir s'il avait un autre âge, éventuellement un autre sexe ou s'il provenait d'un autre centre d'acquisition. Pour ce faire, différentes méthodes de régression (processus gaussien, réseau de neurones...) seront mises en œuvre et comparées.

Ce problème de régression est un défi majeur auquel les chercheurs font face actuellement pour pouvoir utiliser des données de masse dans le cadre de l'imagerie médicale. Des précautions particulières doivent être prises pour résoudre ce problème dans le cas d'images provenant de plusieurs bases de données. En effet, ces dernières ne sont pas obligatoirement appariées en âge, sexe et proportion témoins/patients. Les variables confondantes peuvent ainsi être corrélées avec le diagnostic, conduisant possiblement à une mauvaise interprétation des résultats de classification. Si l'effet des variables confondantes est mal supprimé, le classifieur peut alors utiliser des informations non pertinentes basées sur ces variables pour réaliser la prédiction.

Le cadre applicatif de ce stage concerne l'aide au diagnostic pour la maladie d'Alzheimer à partir de mesures (volume, épaisseur corticale) caractérisant la morphologie de plusieurs structures cérébrales. Ce stage s'appuiera sur une cohorte d'environ 1500 individus (témoin et sujets souffrant de la maladie d'Alzheimer) provenant de différentes bases de données publiques.

*Envoi de candidature (CV+lettre de motivation+ relevés de note des deux ou trois dernières années) à Sylvain Faisan ([faisan@unistra.fr](mailto:faisan@unistra.fr)) et à Vincent Noblet ([vincent.noble@unistra.fr](mailto:vincent.noble@unistra.fr)).*