Cueillette et Traitement de Données

en Neuroscience Cognitive

Projet 1

Description du projet :

Le projet vise à mieux comprendre certains phénomènes liés à la maladie de Alzheimer, notamment des facteurs de risque et des biomarqueurs des syndromes moteurs. Nous allons aussi utiliser le machine learning pour essayer de prédire l’impact de la maladie sur la santé cognitive et mentale.

Description des données :

Notre banque de données (PPMI database) contient les variables suivantes :

PATNO = numéro des patients

Age = âge

Sex = sexe

EDUCYRS = années d’éducation

duration = la durée de la maladie

ageonset = âges des participants lors des premiers symptomes

agediag = âges des participants lors du diagnostic

rigidity = niveau de rigidité musculaire

tremor = niveau de tremblement

upsit = niveau de changement dans l’olfaction

gds = échelle de dépression

hvlt\_retention = échelle cognitive du HVLT

quip = QUIP scape

scopa = échelle du sommeil SCOPA

stai = échelle d’anxiété STAI

moca = échelle de cognition générale MOCA

r\_striatum = volume du striatum droit

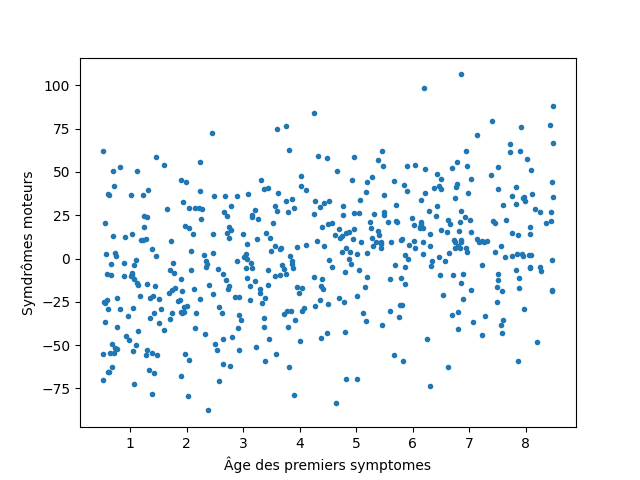
l\_striatum = volume du striatum gauche

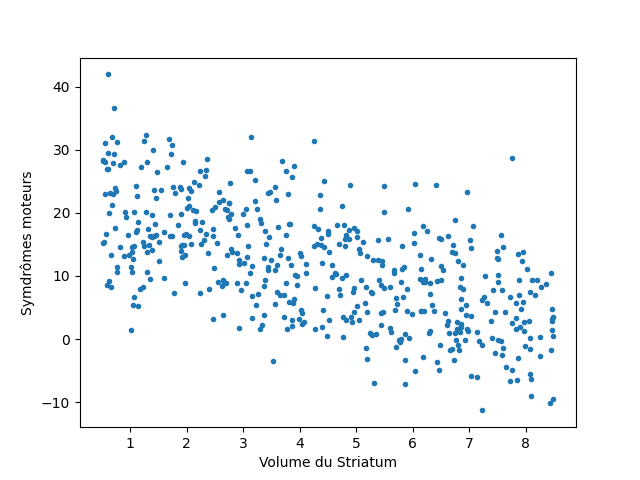
Hypothèses :

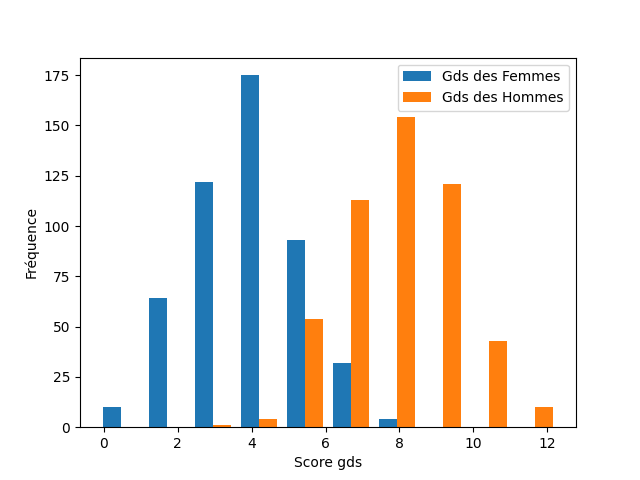
Nous faisons l’hypothèse que :

1. Plus les premiers syndromes se manifestent tôt, moins les syndromes moteurs sont graves.
2. Plus le volume moyen des striatums est grand, moins il y a de syndromes moteurs.
3. Les femmes sont moins dépressives que les hommes.
4. Le score dépressif peut être prédit à partir du reste des variables du jeux de données.
5. Le volume moyen des striatums pourrait prédire un score au MOCA fort ou faible.

Exemples de Graphique pour les hypothèses :







Résultats :

Hypothèse 1 :

Quand l’âge des premiers symptômes augmente d'un an, cela conduit à une aggravation des symptômes moteurs de 0,04 unité, avec un effet significatif (p = 0,014)

Hypothèse 2 :

Quand le volume moyen des striatums augmente d'une unité cela conduit à une aggravation des symptômes moteurs de 1,8 unité, avec un effet significatif (p < 0,000)

Hypothèse 3 :

Il n’y a pas de différence statistiquement significative entre le GDS moyen chez les hommes et chez les femmes (t = -0.125, p = 0,9).

Planning :

Semaine 1 :

Définir la classe et les méthodes de notre analyse (2 heures)

Semaine 2 :

Explorer les meilleurs algorithmes d’apprentissage supervisé pour prédire le groupe du participant (0 : FC faible; 1 : FC élevées) (2 heures)

Semaine 3 :

Explorer les meilleurs algorithmes d’apprentissage non-supervisé pour prédire le score au MOCA d’un participant. (2 heures)

Semaine 4 :

Visualisation de la performance des modèles (2 heures)

Semaine 5 :

Gestion des erreurs (2 heures)

Semaine 6 :

Automatisation du traitement (2 heures)