Hadrien Marcellin
Hector Parmantier

Project Proposal

Object *:* MRI and Alzheimers: Can we diagnose Alzheimer from MRI's numerical features ?

1. Data Description :

The dataset is provided by Kaggle and available here :
https://www.kaggle.com/jboysen/mri-and-alzheimers

The dataset contains nine discriminating features extracted from brain MRI. It is divided in two datasets:

- *cross-sectional MRI:* data comes from 416 subjects aged between 18 and 96. The subjects are all right-handed and include both men and women. 100 of the included subjects over the age of 60 have been clinically diagnosed with very mild to moderate Alzheimer's disease.

- *longitudinal MRI:* data comes from 150 subjects aged from 60 to 96. Here again, the subjects are all right-handed and include both men and women. 72 of the subjects were characterized as nondemented throughout the study whereas 64 of the included subjects were characterized as demented at the time of their initial visits and remained so for subsequent scans.

2. Problem definition

In this project, we wish to identify the features from the MRI images having a strong correlation whether the individual is subject to Alzheimer.

In this perspective we are interested in building a graph based on the features of each subjects. Using the properties of the graph structure, we will try to extract relevant patterns and clusters in the dataset.

The project will apply state of the art graph mining and machine learning techniques to extract these patterns. The different techniques will be compared based on their corresponding performances.

Additionally, with the given dataset, we are freed from the extraction of the RMI features. Therefore, a preprocessing step in extracting relevant features is not necessary. However, a prior exploration of the data will still be necessary in order to clean it (e.g discard potential outliers, missing values handling) and assess the best vectorization method if needed.