

Peer review 14

March 2025

1 Project description

The goal of this project is to investigate if dimension reduction and clustering can be applied to MRI brain scans to predict the severity of alzheimers disease.

There are 400 total observations in the data used, but only 224 usable as some of them dont have any CDR rating.

There are three dimensionality reduction techniques used: Local linear embedding, autoencoders and convolution autoencoders. The two clustering methods used are K-means and DBSCAN. Each dimensionality reduction technique was applied to the front, side and top MRI scans, then followed by clustering.

2 Constructive feedback

- While the data description includes various patient-specific variables, these are not used in the analysis in any way. It may be worth reconsidering whether to include them in the paper at all, or alternatively, exploring how they could be incorporated.
- The Gap statistic could be used to figure out what the optimal amount of clusters are for Kmeans as there might be more distinct clusters.
- You state this in the report: "DBScan is however highly sensitive to the epsilon and minimum points chosen and assumes that clusters have similar density, which may not always be the case due to varying regions of the brain.". HDBSCAN is a method you could try as it does not use epsilon and it handles different density clusters.
- As you mentioned in the paper, you could use a GMM since both the DBSCAN and the Kmeans proved insufficient results.

Organization:

- The report is very well organized in general with clear sections and logical flow.
- Some pictures in the appendix does not fit inside the page. Only half of the network design is visible and the last pictures does not have the whole headline.