UNIVERSITÉ CÔTE D'AZUR

MSc. DSAI

Topological Data Analysis of Human Brain Data

Internship report

June 2022

par

Ufuk Cem Birbiri

Academic tutor: Mathieu Desroches

Inria Center at UCA

Inria

1 What I have done till now

Till now, I have implemented the Mapper algorithm and it is working without a problem. The key elements of the Mapper is parameters such as resolution, gain, filters, clustering algorithm etc. Filters are very important because they map the data points into a vector space where the data points are clustered. When we change the filter we change the Mapper result. The potential filters are :

— **A column in the dataset** : This column will be or main feature where we want to separate data points. For example, 'Age' column can be a filter since we want to cluster the patients according to their age and we want to analyse other features(cognitive, motor, brain surface area, thickness, volume, Gray matter volume, white matter volume etc.).

— L**-infinity centrality** ,

— L_p**- Norm**,

— **Singular Value Decomposition (SVD)**,

— **PCA**

Selecting appropriate parameters is not a problem because I implemented a code that searches the best parameters and visualizes the result. So I can pick them easily.
The Mapper result really changes according to dataset. It is very important that which columns we use in the Mapper. Here, the question is : What are we searching for ? What is our research question ?
Before asking this question, I want to write a quick recap about our meeting with Mathieu and Serafim.

2 A recap for the last meeting

I needed to write my internship report before August 15, so I used my previous Mapper result. In this result, I used Gender column and L-Infinity centrality as filters, and cognition, emotion and motor columns as data columns. I found 4 different communities : 1 group of fully males, 3 group of fully females. There is no homogeneity in the groups based on age, since I used Gender column as the first filter. This filter divide the data into males and females, and the L-Infinity centrality found subgroups.
After feature analysis of these groups I found, for example :
The male group have problems of sadness and loneliness. Likewise, The female groups perform better at language, and cognitive tasks than patients within the male group. Ac-

tually this analysis is more detailed, but I want to give you an idea about my internship report. This part is not that important so let's move on.

In the meeting we decided that Gender column is meaningless because what we want to analyse is the Age of patients with cognitive, emotion, or motor skills. In our data, Age has 4 categories :[22-25], [26-30], [31-35] and, [+36]. Also Age can be related to many things in the dataset such as brain surface area, thickness, brain volume, Gray matter volume, white matter volume etc. If we can relate the 'Age' with brain thickness or surface area or Gray matter volume, it will be a more powerful research and we can write a more powerful paper. So at this point we need to overview the dataset and decide what kind of questions we should ask for this research. The potential questions are :

— Is the Gray matter volume decreases when people gets old ?
— What about white matter volume ?
— Does brain thickness reduces in older people compared to young ones ?
— What about the surface area or brain volume ?
— The cognitive skills are getting worse when people gets old. Is it related to Gray matter volume in the brain ?
— ...

We can ask many questions. Here, what we need is a good literature review (with your help) and a good question to ask(maybe with the help of a neuro-scientist). And this question should be specific because in the whole data there are many many columns for each section :

Gray matter volume(6 columns) :
1. Left hemisphere cortical Gray matter volume,
2. Right hemisphere cortical Gray matter volume,
3. Total cortical Gray matter volume,
4. Total subcortical Gray matter volume,
5. Total Gray matter volume,
6. Supratentorial volum.

White matter volume(3 columns) :
1. Left hemisphere cortical white matter volume,
2. Right hemisphere cortical white matter volume,
3. Total cortical white matter volume.

Volume (Subcortical) Segmentation (45 colums) :

1. Left-Lateral-Ventricle Volume,
2. Left-Inf-Lat-Vent Volume,
3. Left-Cerebellum-White-Matter Volume,
4. Left-Cerebellum-Cortex Volume,

...

Surface Area (68 columns)

1. Left bankssts Surface Area,
2. Left caudalanteriorcingulate Surface Area,
3. Left caudalmiddlefrontal Surface Area,
4. Left cuneus Surface Area,

...

Surface Thickness (68 columns)

1. Left bankssts Average Thickness,
2. Left caudalanteriorcingulate Average Thickness,
3. Left caudalmiddlefrontal Average Thickness,
4. Left cuneus Average Thickness,

...

Sensory (5 columns)

1. Audition (Words in Noise),
2. Olfaction (Odor Identification Test),
3. Pain (Pain Intensity and Interference Surveys),
4. Taste (Taste Intensity Test),
5. Contrast Sensitivity (Mars Contrast Sensitivity).

So if we are going to choose to analyse brain thickness, surface area, or volume, what specific brain region are we working on ? There are so many brain regions in the dataset and I cannot put them all together, it will be a mess. We should decide which columns to pick. Personally I have no idea. However, I made some literature review and found some articles about age vs gray matter volume, age vs cognition, or age vs thickness. And I implemented the Mapper algorithm according to these papers results and compared my result with theirs. If I will find similar results compared to papers in literature, then it would be awesome. Let's see the papers in literature first :

3 Literature Review

— ***Global associations between regional Gray matter volume and diverse complex cognitive functions : evidence from a large sample study (2017) [1]*** :

"Better performance scores on all tasks except non-verbal reasoning were associated with greater regional Gray matter volume across widespread brain areas."

There no 'aging vs cognitive skills' in this paper. The patient's average age is 20.8 years. They say, *when Gray matter volume increases, the cognitive scores increase.*

— ***Gray Matter Volume and Cognitive Performance During Normal Aging. A Voxel-Based Morphometry Study(2018) [2]*** :

"Our results show smaller bilateral Gray matter volume in the older adults(mean age 71 years) relative to the middle-age group(mean age 41 years), in several cerebral and right cerebellar regions involved in language and executive functions. Importantly, our results also revealed smaller Gray matter volume in the right cerebellum in older adults relative to middle-age, supporting the idea of a complex cognitive role for this structure."

As I understand, Gray matter volume decreases over age, and cognitive test scores also decrease.(This is a general interpretation but of course we have to consider cerebellum or cerebral regions).

— ***Mapping cortical change across the human life span(2003) [3]*** :

"A nonlinear decline in Gray Matter Density with age, which was most rapid between 7 and about 60 years, over dorsal frontal and parietal association cortices on both the lateral and interhemispheric surfaces."

The Gray matter density decline when people get old.

— ***Age Differences in Prefrontal Surface Area and Thickness in Middle Aged to Older Adults(2016) [4]*** :

"We investigated age differences in prefrontal surface area and cortical thickness in healthy adults between the ages of 51 and 81 years. We found that older age was associated with smaller surface area in the dorsolateral prefrontal and orbitofrontal cortices but greater cortical thickness in the dorsolateral prefrontal and anterior cingulate cortices."

So the paper says :

With aging, Smaller surface area in : ~~dorsolateral prefrontal~~(not in our dataset) and in orbitofrontal cortices.–> In our dataset there are these columns for orbitofrontal surface area :

Left lateral-orbitofrontal Surface Area.✓

Left medial-orbitofrontal Surface Area.✓

Right lateral-orbitofrontal Surface Area.✓

Right medial-orbitofrontal Surface Area.✓

I put them in the final dataset for my Mapper model.

And the paper says :Greater cortical thickness in ~~dorsolateral prefrontal~~ and in anterior cingulate cortices. –> In our dataset there are these columns for anterior cingulate cortice thickness :

Left caudal-anterior-cingulate Average Thickness.✓

Left rostral-anterior-cingulate Average Thickness.✓

Right caudal-anterior-cingulate Average Thickness.✓

Right rostral-anterior-cingulate Average Thickness.✓

I put them in the final dataset for my Mapper model.

4 My final dataset(998,50)

My final dataset consists of these features :

1. Cognition columns
2. Motor columns
3. Emotion columns
4. Age and Gender columns
5.

In Anterior Cingulate Cortex :

- Left caudal-anterior-cingulate Surface Area,
- Left rostral-anterior-cingulate Surface Area,
- Right caudal-anterior-cingulate Surface Area,
- Right rostral-anterior-cingulate Surface Area,

- Left caudal-anterior-cingulate Average Thickness,
- Left rostral-anterior-cingulate Average Thickness,
- Right caudal-anterior-cingulate Average Thickness,

- Right rostral-anterior-cingulate Average Thickness,

In Orbitofrontal Cortex :

- Left lateral-orbitofrontal Surface Area,
- Left medial-orbitofrontal Surface Area,
- Right lateral-orbitofrontal Surface Area,
- Right medial-orbitofrontal Surface Area,

- Left lateral-orbitofrontal Average Thickness,
- Left medial-orbitofrontal Average Thickness,
- Right lateral-orbitofrontal Average Thickness,
- Right medial-orbitofrontal Average Thickness

6. White matter volume :

- Total cortical Gray matter volume.
- Total Gray matter volume.

7. **Supratentorial volume**

8. **Total cortical white matter volume**

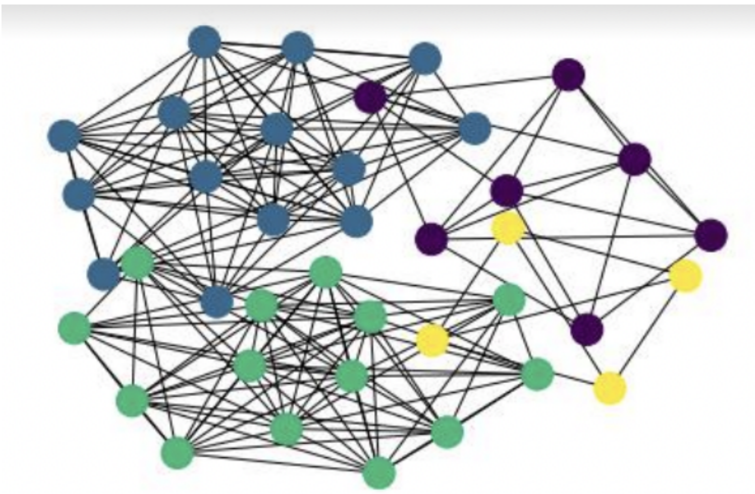

FIGURE 1 – *The Mapper result colored by Age*

5 Results

I achieved 4 different communities. It is not surprising because we have 4 different age categories in the data : [22-25], [26-30], [31-35], and [+36]. Since I used the Age column as the first filter we have four categories. The Mapper result can be shown in Fig.1. And the 4 different communities can be seen in Fig.2.

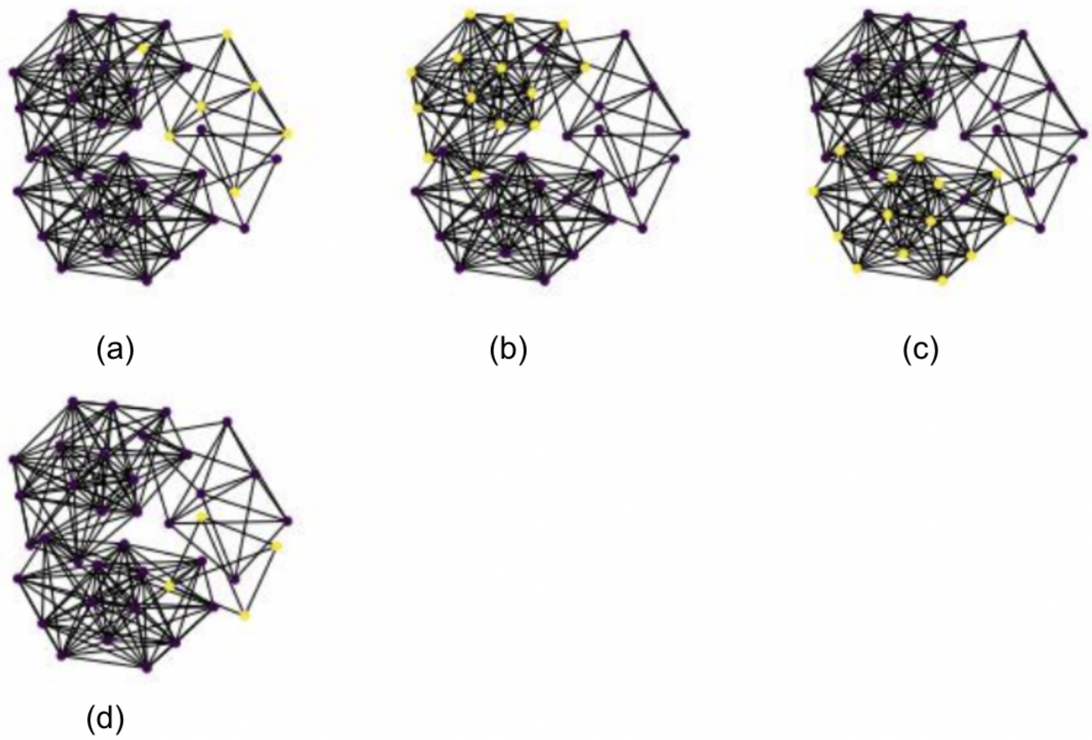

FIGURE 2 – *The communities in the Mapper result. (a) is the first community called C1 :[26-30]. (b) is C2 :[31-35]. (c) is C3 :[22-25]. (d) is C4 :[+36].*

I have analyzed the cognitive, emotional, motor, and Gray matter, white matter, thickness and surface area in each group, and identified which features make these groups different. Clinical features specific to groups are listed below where p-values are in ascending order from top to bottom i.e. first row in the tables is the most representative feature of that community. I used different colors while investigating the important features for communities to make it more readable.

Total Gray Matter Volume [2] says that Gray matter volume decreases over age, and cognitive test scores also decrease. And also [3] says that the Gray matter density decline when people get old. So let's look at our results. The *Total Gray Matter Volume* is a

Clinical features	C_1[26-30]	C_2[31-35]	C_3[22-25]	C_4[+36]
Age	✓	✓	✓	✓
Language/Vocabulary ComprehensionGray	✓	✓		
Total Gray Matter Volume	✓	✓	✓	
Total Cortical Gray Matter Volume	✓	✓	✓	
Total White Matter Vol	✓			
Supratentorial volume	✓	✓	✓	
Right caudal-anterior-cingulate Surface Area	✓	✓		
Fluid Intelligence	✓	✓		

TABLE 1 – **(A) Clinical features for C1 :[26-30]**.

significant feature for C1[26-30], C2[31-35] and C3[22-25]. The mean values of *Total Gray Matter Volume* are :
C3[22-25] -> 706725.3
C1[26-30] -> 693064.8
C2[31-35] -> 668933.9
(C4[+36] -> 644776.6). The *Total Gray Matter Volumes* are decreasing when age increase. We can conclude that *Total Gray Matter Volume* is a representative feature for C1, C2 and C3 (not C4) and it reduces by age. These findings are parallel with the papers [2, 3].

Total cortical gray matter volume The *Total cortical gray matter Volume* is a significant feature for C1[26-30], C2[31-35] and C3[22-25]. The mean values of *Total cortical gray matter Volume* are :
C3[22-25] -> 526921.5
C1[26-30] -> 515001.4
C2[31-35] -> 496823.7
(C4[+36] -> 474106.6).
Also, the total cortical gray matter volume is decreasing when people get old. It is a representative feature for for C1, C2 and C3 (not C4). These result also supports the papers [2, 3].

Cognitive Scores [1] declines that when Gray matter volume increases, the cognitive scores increase. Let's see. Fluid Intelligence appeared as a discriminative feature in C1[26-30] and, C2[31-35]. The mean values of *Fluid Intelligence* and Total Gray Matter Volume

Clinical features	C_1[26-30]	C_2[31-35]	C_3[22-25]	C_4[+36]
Age	✓	✓	✓	✓
Total Gray Matter Volume	✓	✓	✓	
Total Cortical Gray Matter Volume	✓	✓	✓	
Right lateral-orbitofrontal Surface Area		✓	✓	
Left lateral-orbitofrontal Surface Area		✓	✓	
Left lateral-orbitofrontal Avg Thickness		✓	✓	✓
Fluid Intelligence	✓	✓		
Gender		✓	✓	
Supratentorial volume	✓	✓	✓	
Right lateral-orbitofrontal Avg Thickness		✓	✓	
Left rostral-anterior-cingulate Surface Area		✓	✓	
Right caudal-anterior-cingulate Surface Area	✓	✓		
Language/Vocabulary Comprehension	✓	✓		
Right medial-orbitofrontal Avg Thickness		✓	✓	
Spatial Orientation		✓		
Right medial-orbitofrontal Surface Area		✓		
Negative Affect-Anger		✓	✓	
Working Memory		✓		
Locomotion		✓		✓
Language/Reading Decoding		✓		
Left medial-orbitofrontal Surface Area		✓		
Right rostral-anterior-cingulate Surface Area		✓		
Episodic Memory		✓		

TABLE 2 – **(A) Clinical features for C2 :[31-35]**.

are shown below, respectively :
C1[26-30] -> 17.46 - 693064.8
C2[31-35] -> 16.22 - 668933.9.
Over age, fluid intelligence score decreases as well as total gray matter volume. So this result support [1].

What about Language/Vocabulary Comprehension ? Language/vocabulary comprehen-

Clinical features	$\mathbf{C_1}$[26-30]	$\mathbf{C_2}$[31-35]	$\mathbf{C_3}$[22-25]	$\mathbf{C_4}$[+36]
Age	✓	✓	✓	✓
Gender		✓	✓	
Total Gray Matter Volume	✓	✓	✓	
Total Cortical Gray Matter Volume	✓	✓	✓	
Left lateral-orbitofrontal Avg Thickness		✓	✓	✓
Right lateral-orbitofrontal Surface Area		✓	✓	
Negative Affect-Anger		✓	✓	
Left lateral-orbitofrontal Surface Area		✓	✓	
Sustained Attention spec			✓	
Right lateral-orbitofrontal Avg Thickness		✓	✓	
Right caudal-anterior-cingulate Aveg Thickness			✓	
Strength			✓	
Processing Speed		✓		
Left rostral-anterior-cingulate Surface Area		✓	✓	
Sustained Attention sensivity			✓	
Right medial-orbitofrontal Average Thickness		✓	✓	
Supratentorial volume	✓	✓	✓	

TABLE 3 – **(A) Clinical features for C3 :[22-25]**.

Clinical features	$\mathbf{C_1}$[26-30]	$\mathbf{C_2}$[31-35]	$\mathbf{C_3}$[22-25]	$\mathbf{C_4}$[+36]
Age	✓	✓	✓	✓
Right rostral-anterior-cingulate Avg Thickness				✓
Left lateral-orbitofrontal Avg Thickness		✓	✓	✓
Locomotion		✓		✓

TABLE 4 – **(A) Clinical features for C4 :[+36]**.

sion is appeared in again C1[26-30] and, C2[31-35]. The mean values of *Language/Vocabulary Comprehension* and Total Gray Matter Volume are shown below, respectively :

The mean values of *Fluid Intelligence* are :

C1[26-30] -> 111.63 - 693064.8

C2[31-35] -> 107.79 - 668933.9.

Again, when people get old, language/vocabulary comprehension score decreases, as well as total grey matter volume. This result supports [1].

There are other cognitive scores in the tables but there is no other feature that appears in more than one community. So, I only analyse Language/Vocabulary Comprehension and Fluid Intelligence.

Surface Area in Orbitofrontal Cortices [4] claims that with aging, there is a smaller surface area in stdorsolateral prefrontal(not in our dataset) and in orbitofrontal cortices. Let's analyse our results. We have two orbitofrontal columns to support this idea : Left lateral-orbitofrontal Surface Area and, Right lateral-orbitofrontal Surface Area. Let's see their average values in communities :
Right lateral-orbitofrontal Surface Area appeares in C2[31-35] and, C3[22-25]. Their avg values are :

C3[22-25] -> 2671.4
C2[31-35] -> 2522.8
As we can see, the right lateral-orbitofrontal surface area decreases when age increase, which supports [4]. Let's see left one :
Left lateral-orbitofrontal Surface Area
This one also appears in C2[31-35] and, C3[22-25] as a representative feature of the communities. Their avg values are :
C3[22-25] -> 2724.4
C2[31-35] -> 2592.5
We can see that left lateral-orbitofrontal surface area also decreases when people get old. This is also supports [4]. Good!

Avg Thickness in Anterior Cingulate cortices This paper([4]) also says that there is a greater cortical thickness in : ~~dorsolateral prefrontal~~ and anterior cingulate cortices. I added the anterior cingulate cortices columns to the final dataset but there is no anterior cingulate feature apperead as a significant representative feature among communities. So we don't have any resukts to support this idea.

BONUS : Avg Thickness in Left Lateral OrbitoFrontal According to our results, the thickness of Left Lateral OrbitoFrontal is decreasing when age is increasing. There is no paper that supports this idea (I couldn't find, and I didn't search). The *Left Lateral Or-*

bitoFrontal appears in C2[31-35], C3[22-25], C4[+36] with the avg values :
C3[22-25] -> 2.84
C2[31-35] -> 2.79
C4[+36] -> 2.72
As we can see, the thickness is decreasing while age is increasing.

6 Discussion

This report aims to show you what kind of paper we can produce. There are some questions we should ask such as which columns we should choose or what are we analysing ?(thickness, surface area, gray matter, cognition...). Waiting your help :)

Bibliographie

[1] Hikaru Takeuchi, Yasuyuki Taki, Rui Nouchi, Ryoichi Yokoyama, Yuka Kotozaki, Seishu Nakagawa, Atsushi Sekiguchi, Kunio Iizuka, Yuki Yamamoto, Sugiko Hanawa, et al. Global associations between regional gray matter volume and diverse complex cognitive functions : evidence from a large sample study. *Scientific Reports*, 7(1) :1–16, 2017.

[2] Stephen Ramanoël, Elena Hoyau, Louise Kauffmann, Félix Renard, Cédric Pichat, Naïla Boudiaf, Alexandre Krainik, Assia Jaillard, and Monica Baciu. Gray matter volume and cognitive performance during normal aging. a voxel-based morphometry study. *Frontiers in Aging Neuroscience*, 10 :235, 2018.

[3] Elizabeth R Sowell, Bradley S Peterson, Paul M Thompson, Suzanne E Welcome, Amy L Henkenius, and Arthur W Toga. Mapping cortical change across the human life span. *Nature neuroscience*, 6(3) :309–315, 2003.

[4] Vonetta M Dotson, Sarah M Szymkowicz, Christopher N Sozda, Joshua W Kirton, Mackenzie L Green, Andrew O'Shea, Molly E McLaren, Stephen D Anton, Todd M Manini, and Adam J Woods. Age differences in prefrontal surface area and thickness in middle aged to older adults. *Frontiers in aging neuroscience*, 7 :250, 2016.