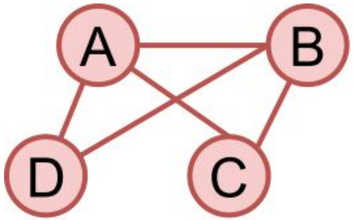
Revealing integrated and segregated structures using a graph neural network

Anna Boronina

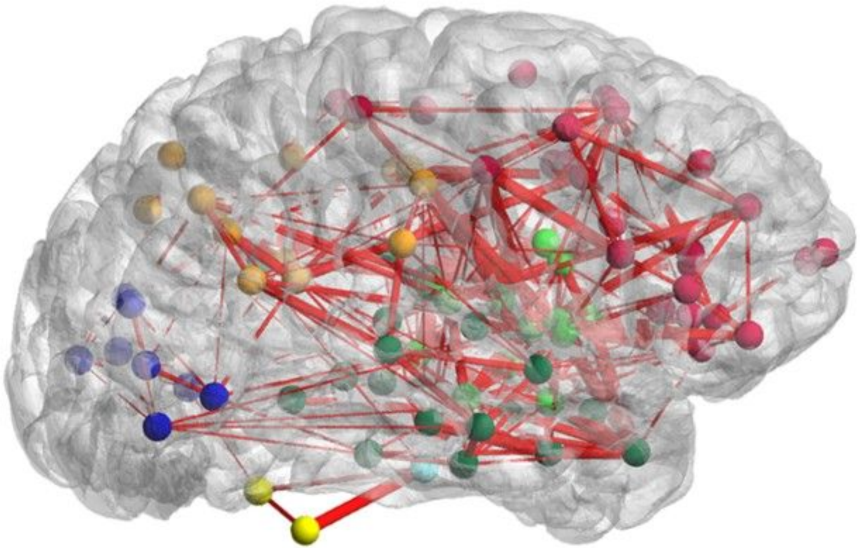
INNOPOLIS UNIVERSITY

Structure

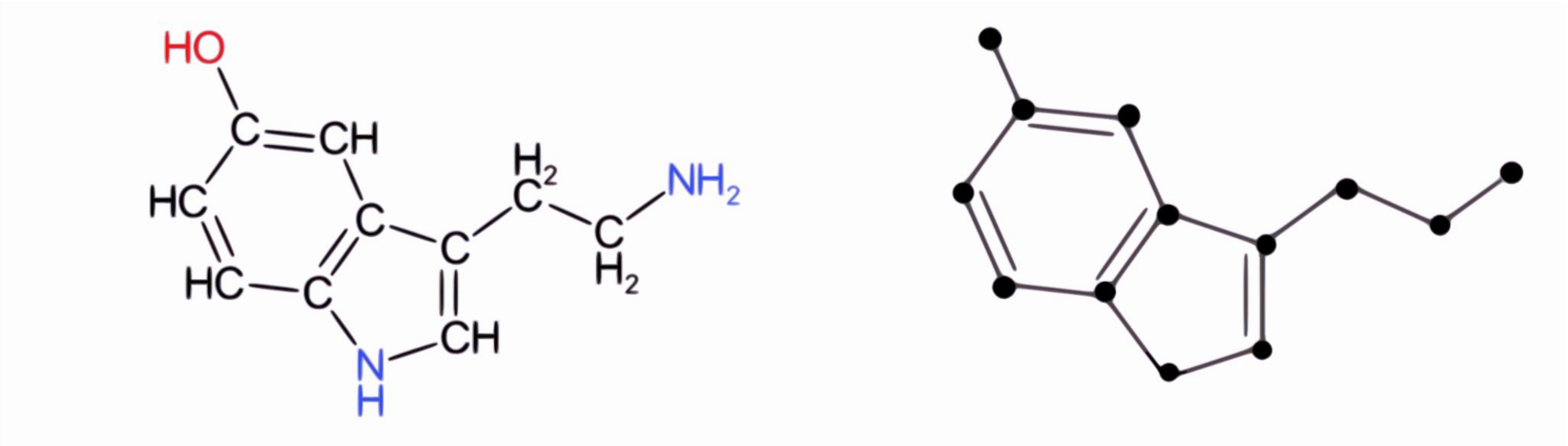
- What are graphs and Graph Neural Networks
- Research gap
- Datasets creation
- GNN architecture
- Results



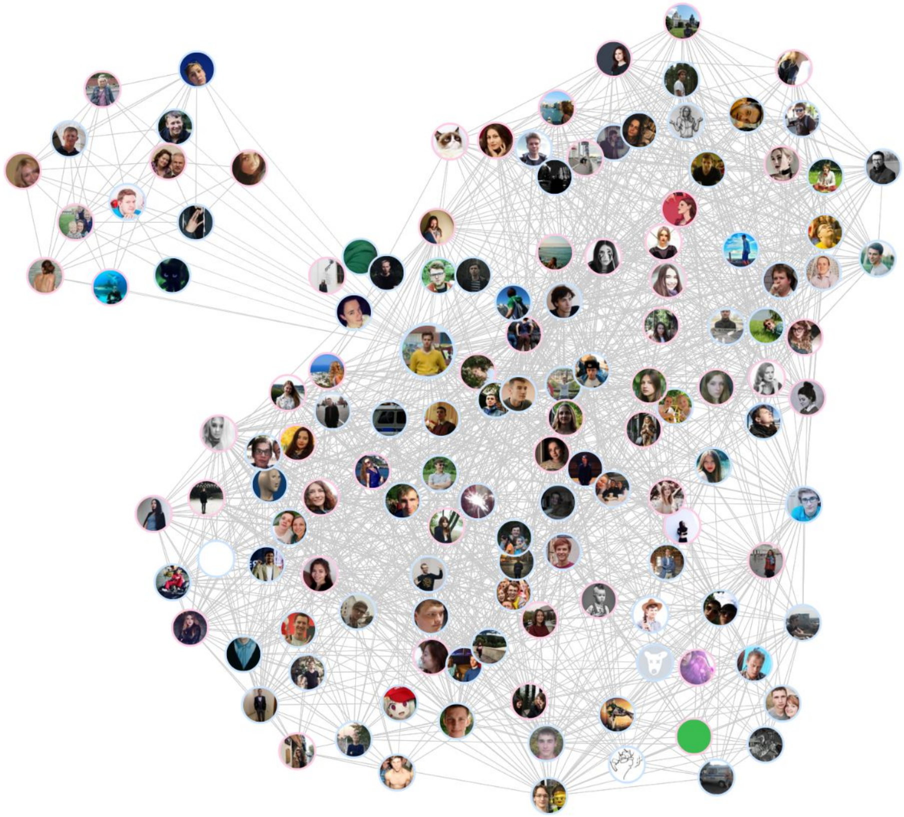
Graphs in real life



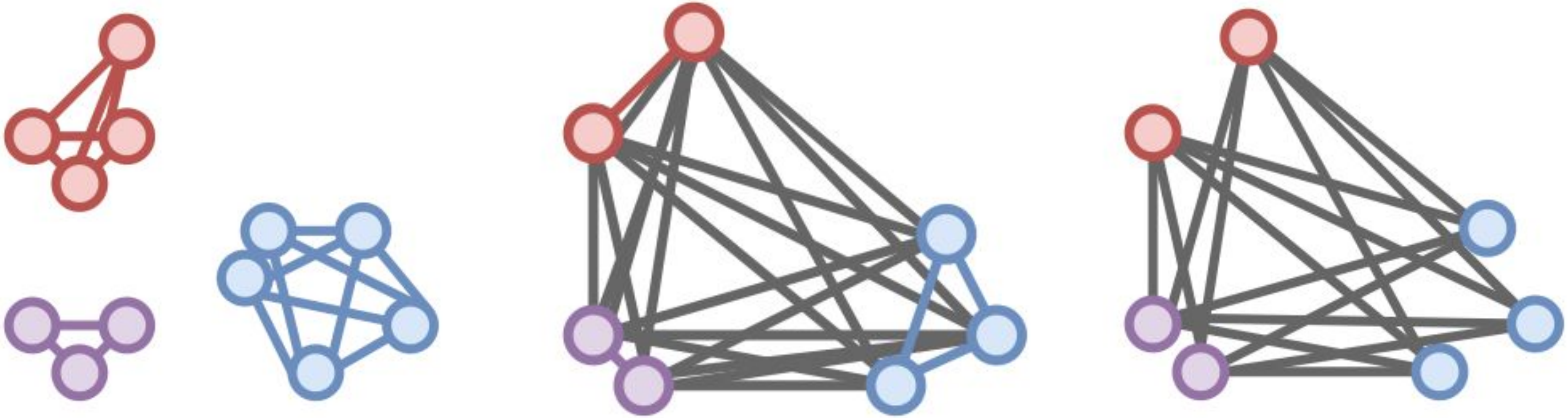
Graphs in real life



Graphs in real life



Segregation and integration



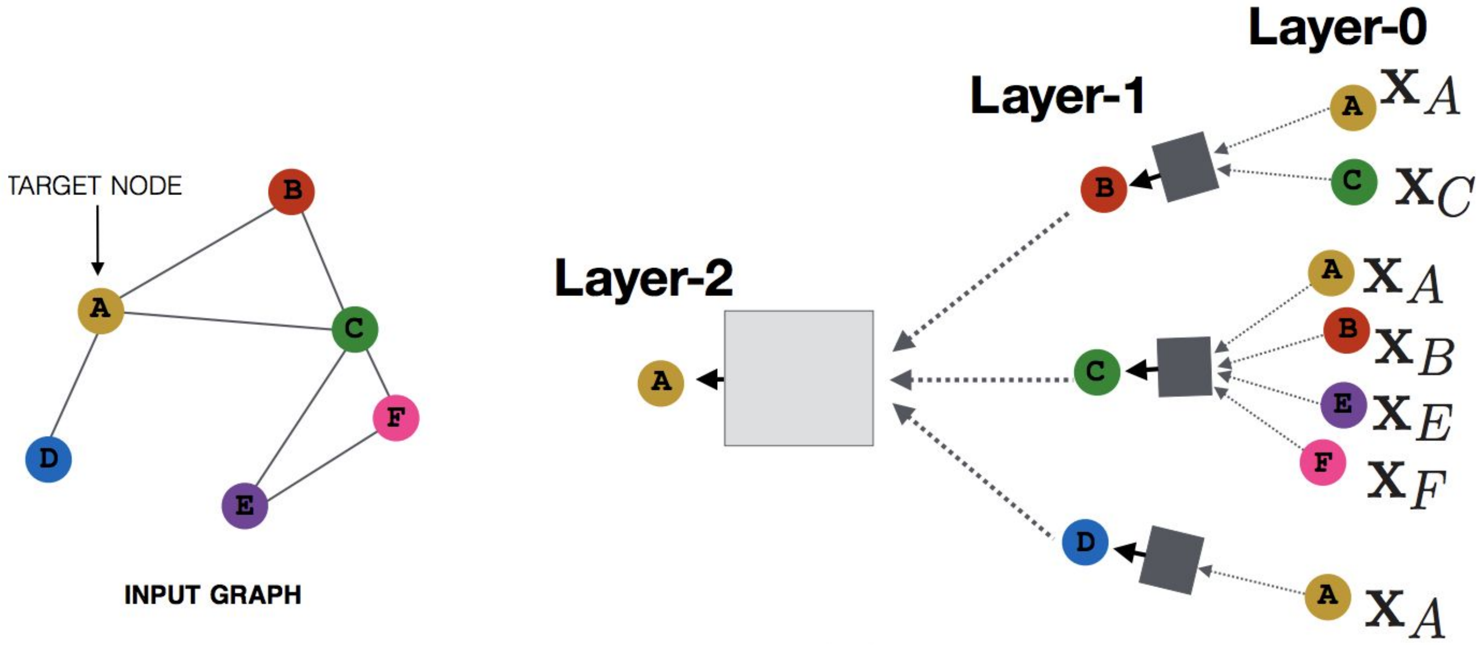
Fully segregated | Neutral | Fully integrated

The balance is important

- **segregation** affects **motor ability** distant brain regions
- **segregation** is important for a **security network**

- **stronger integration** in a brain network implies **better cognitive ability**
- **integration** is important for **city transport system**

Graph Neural Networks



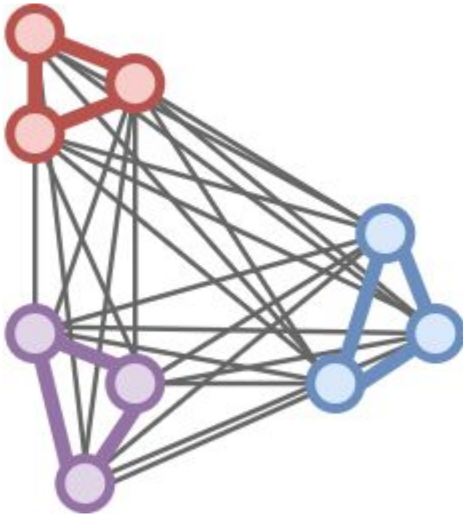
http://snap.stanford.edu/proj/embeddings-www/

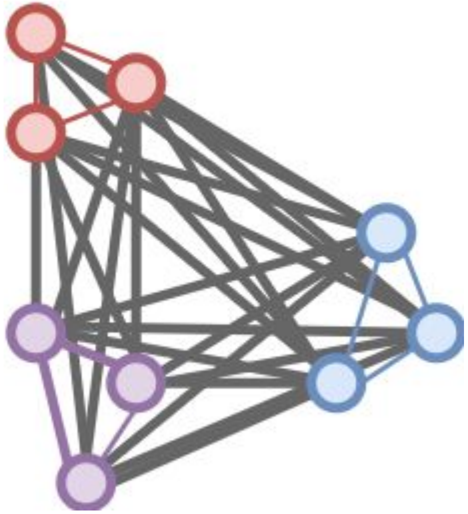
Research Gap

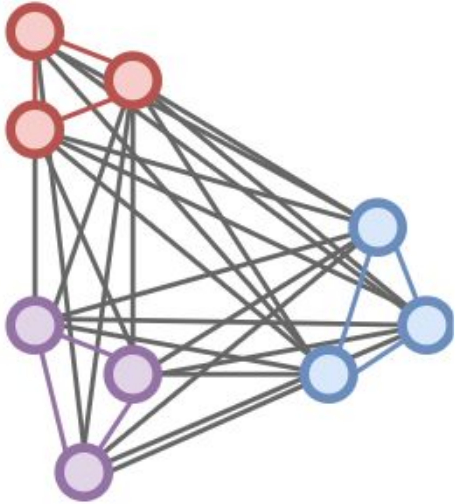
GNNs are mostly applied to real data and rarely interpreted

Data modeling allows to control the input and see what parameters play what role

Segregation Integration Neutral



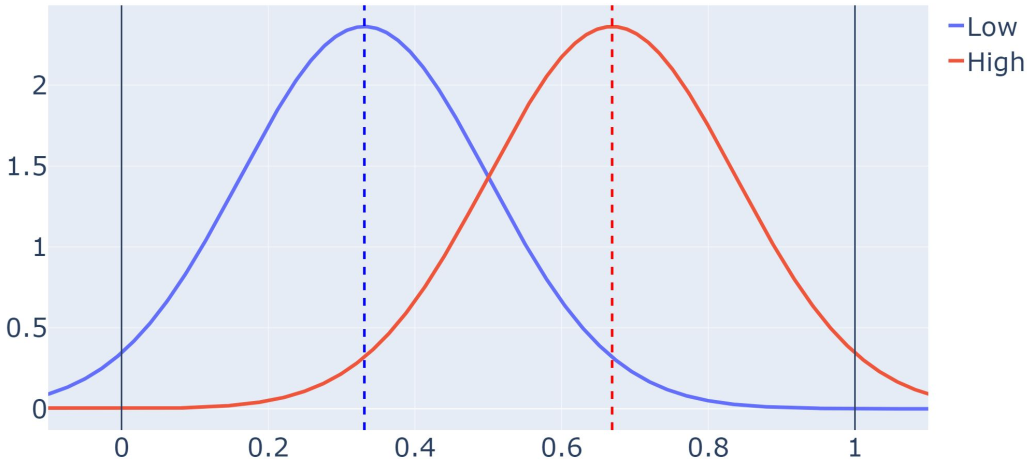




OK, so...

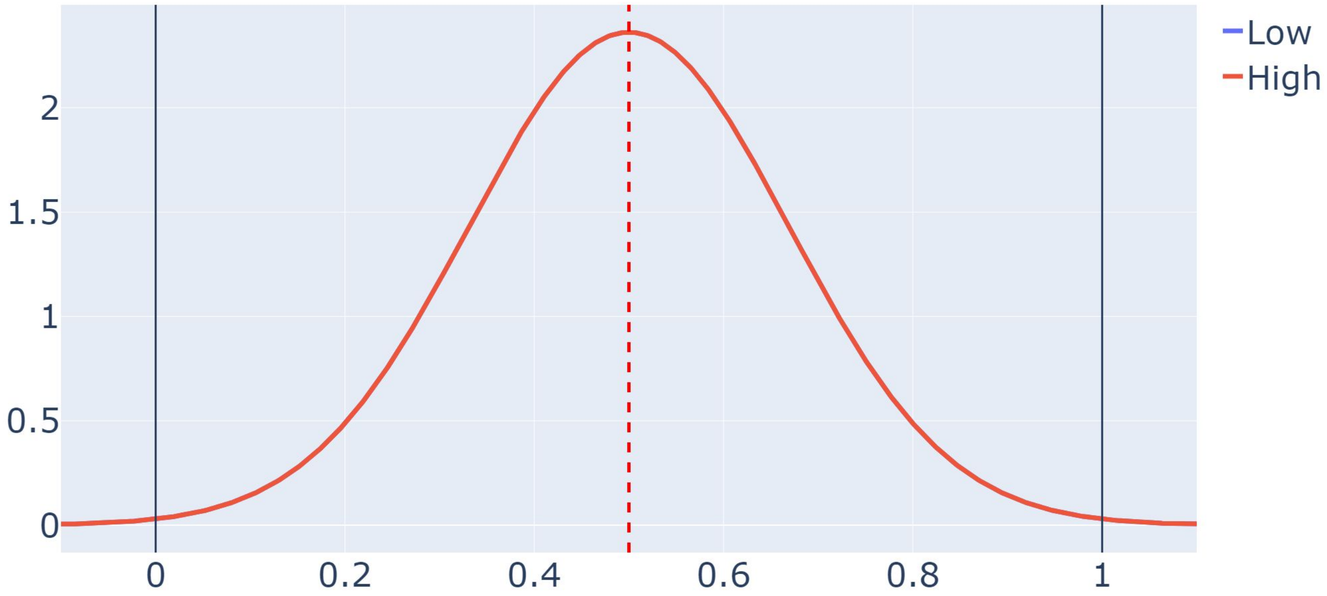
There are two types of connections: **weak** | **strong**

How do we create these connections?



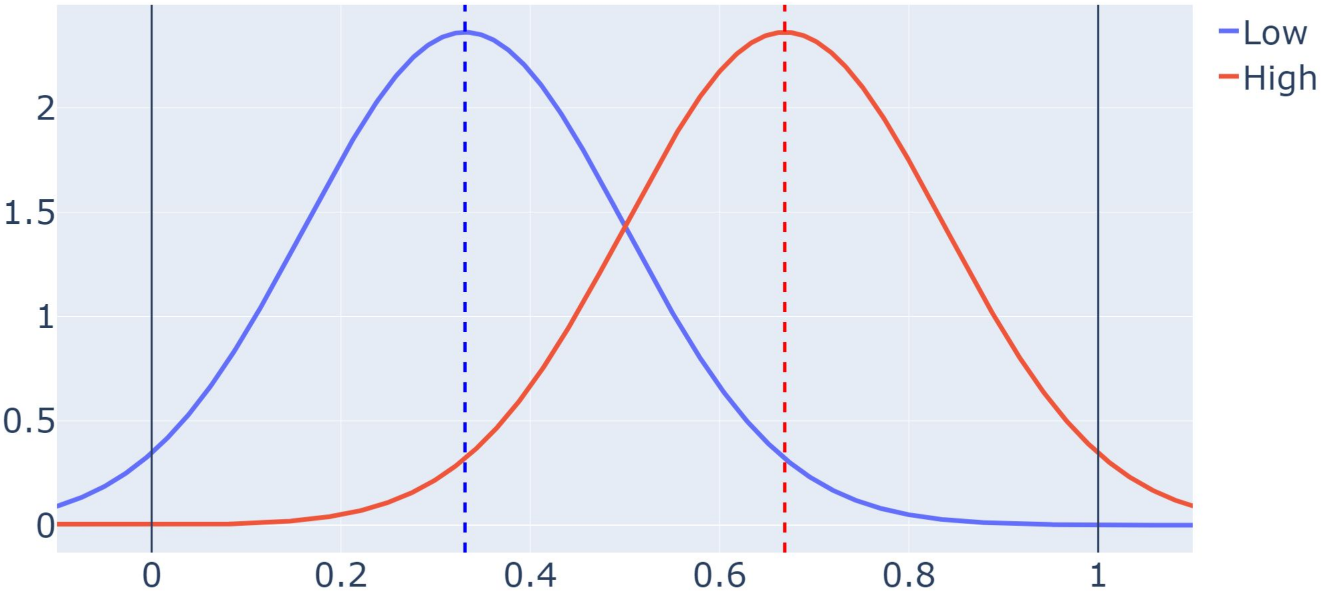
How far L and H distributions should be?

Maybe the same?

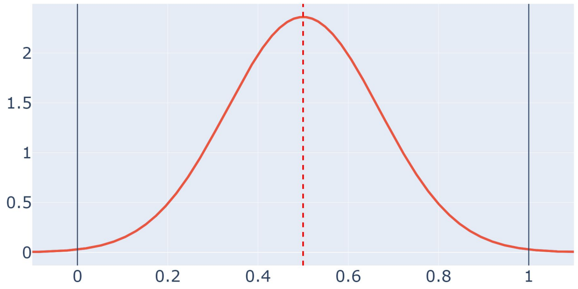


How far L and H distributions should be?

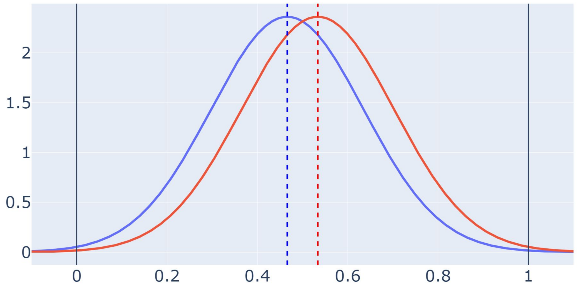
Maybe far away?



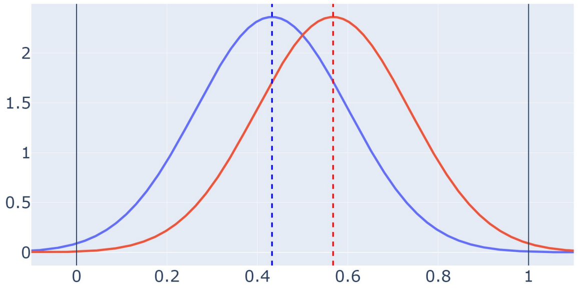
Let’s make it a parameter!



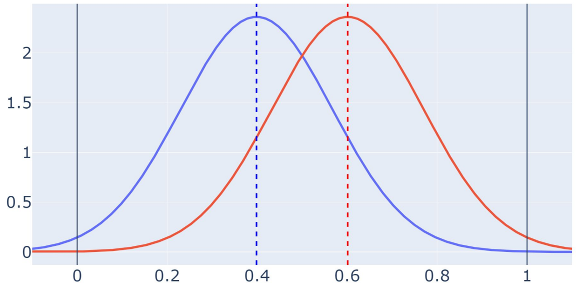
n = 0



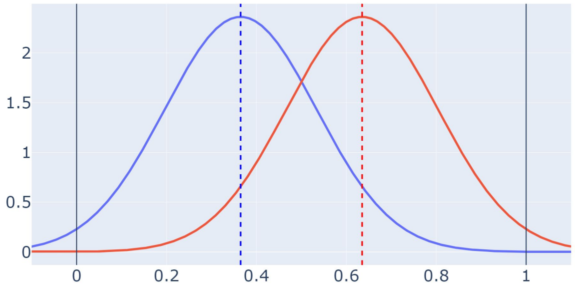
n = 1



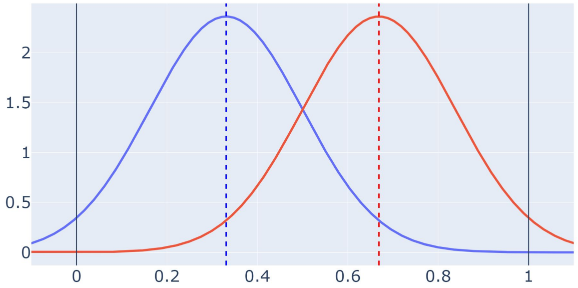
n = 2



n = 3

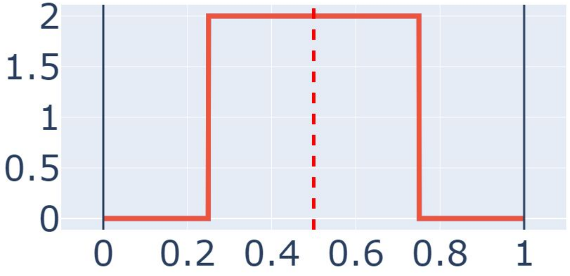


n = 4

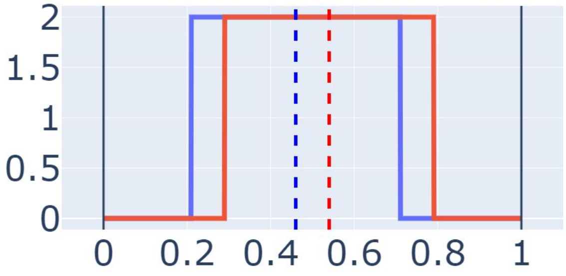


n = 5

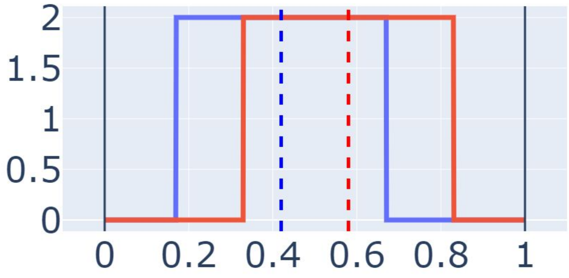
What if we try another distribution?



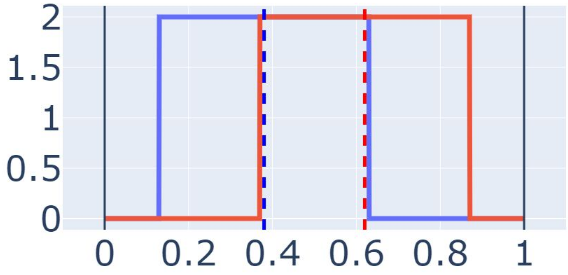
n = 0



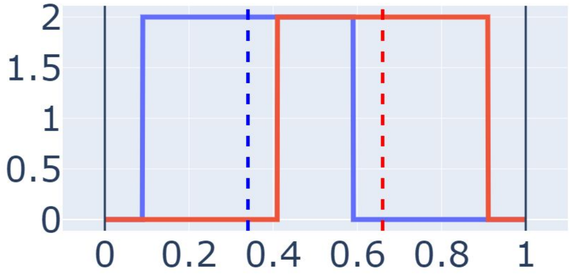
n = 1



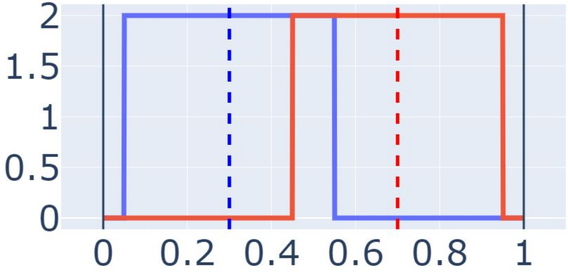
n = 2



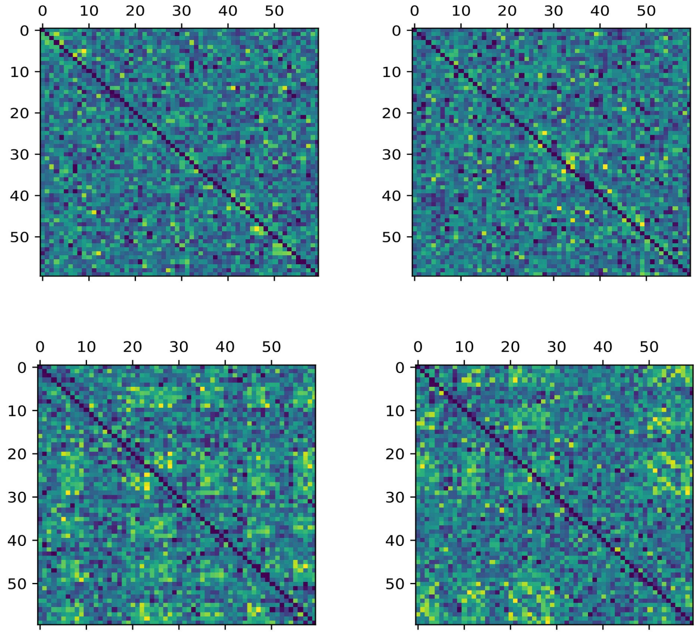
n = 3



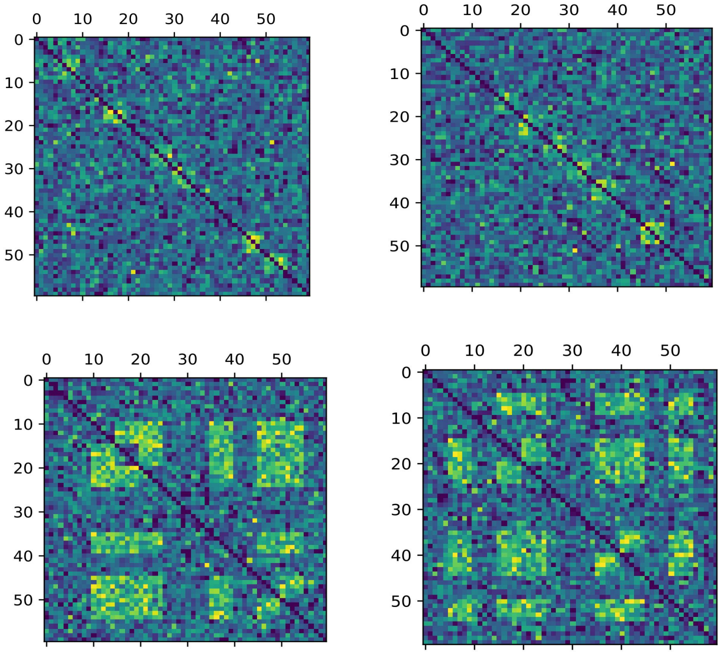
n = 4



n = 5



n = 3



n = 5

OK, so...

Two types of connections: **weak** | **strong**

Distributions for connections: **Normal** | **Uniform**

Three types of graphs: **segregated** | **integrated** | **neutral**

Before finally modeling the graphs

How to choose…

- **number of regions**?
- **number of nodes per region**?

Modularity and Participation

These metrics **reflect degree of segregation and integration** considering **regional information**

$$Q_i = \frac{1}{I}\Sigma_{i,j\in\mathbb{N}}\left[w_{ij} - \frac{k_i k_j}{I}\right]\boxed{\delta_{r_i,r_j}}$$

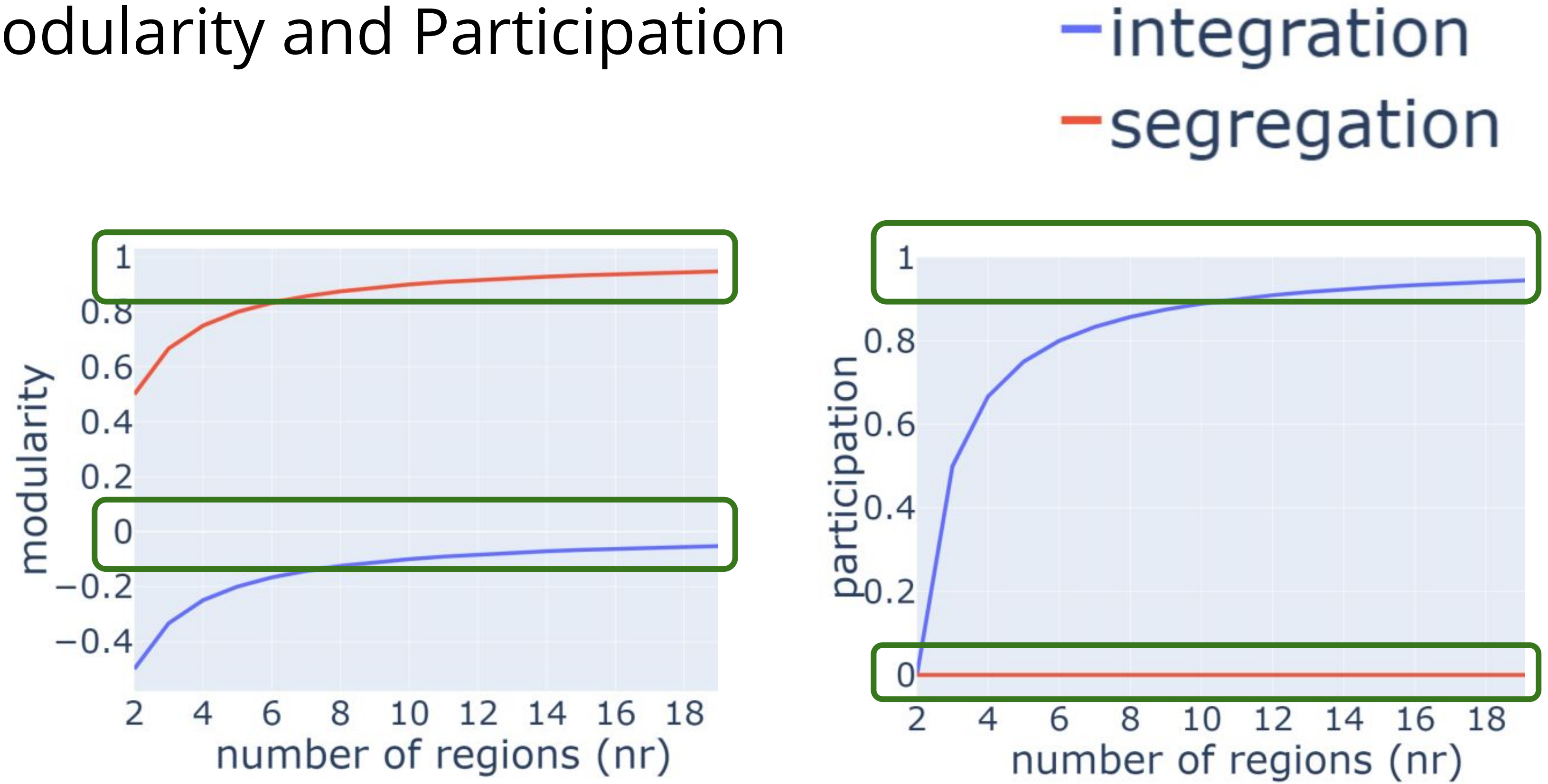
$$\delta_{r_i,r_j} = \begin{cases} 1, & \text{if nodes i and j are in the same region} \\ 0, & \text{otherwise} \end{cases}$$

$$PC_i = 1 - \Sigma_{r=1}^{\boxed{R}}\left(\boxed{\frac{k_i(r)}{k_i}}^2\right)$$

$k_i(r)$ - connectivity strength of node i within region r

R - region number

Modularity and Participation



The final choice

Number of regions: 12

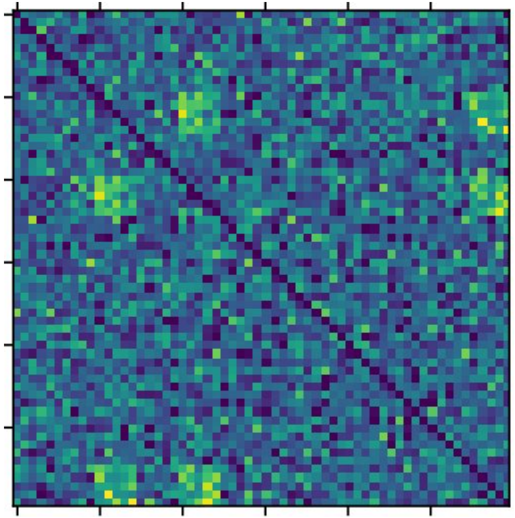
Number of nodes per region: 5

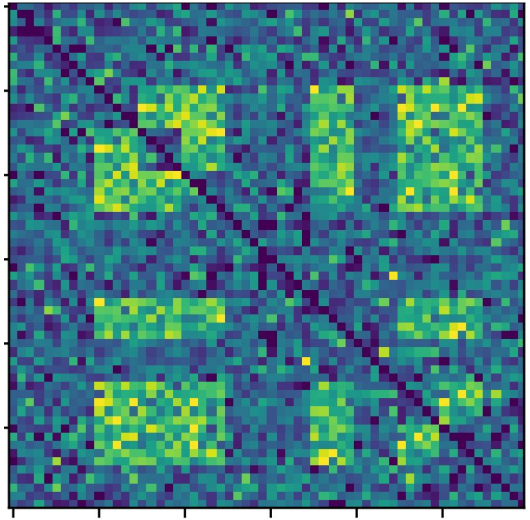
The final choice

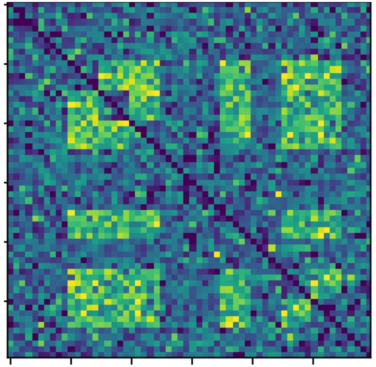
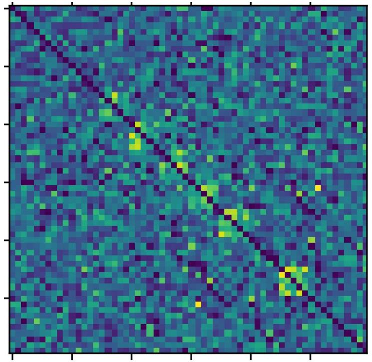
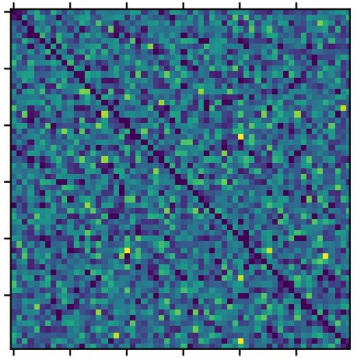
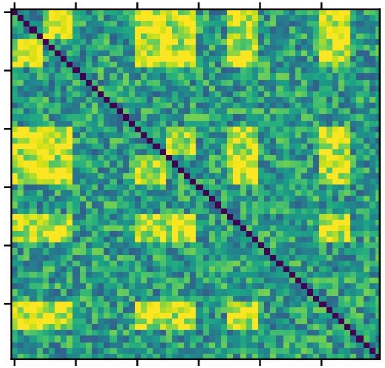
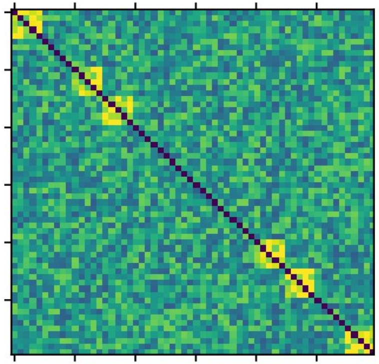
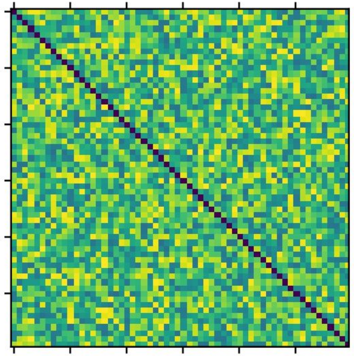
Number of regions: 12

How many are **active**?
3 | 6 | 9 | all

Number of nodes per region: 5

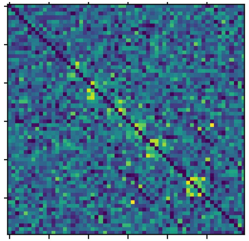




	Integration	Segregation	Neutral
Normal			
Uniform			

❔ Research Questions

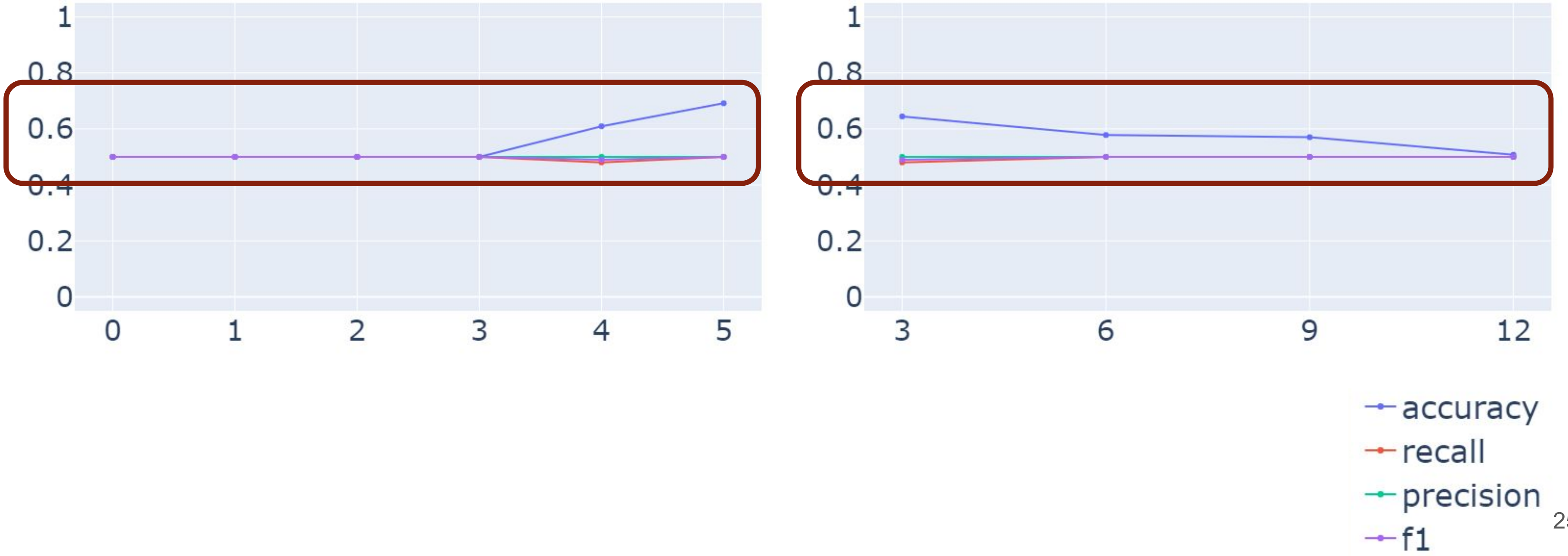
- Will a GNN show **better results** for either segregation or integration?
- Is there any difference depending on the **difference between weak and strong connections** in a graph?
- Does a **type of distribution** play a role?

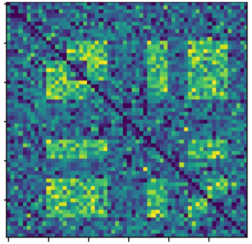


Segregation based on **Normal** distribution

metrics depending on *n*

metrics depending on *nar* for n > 1

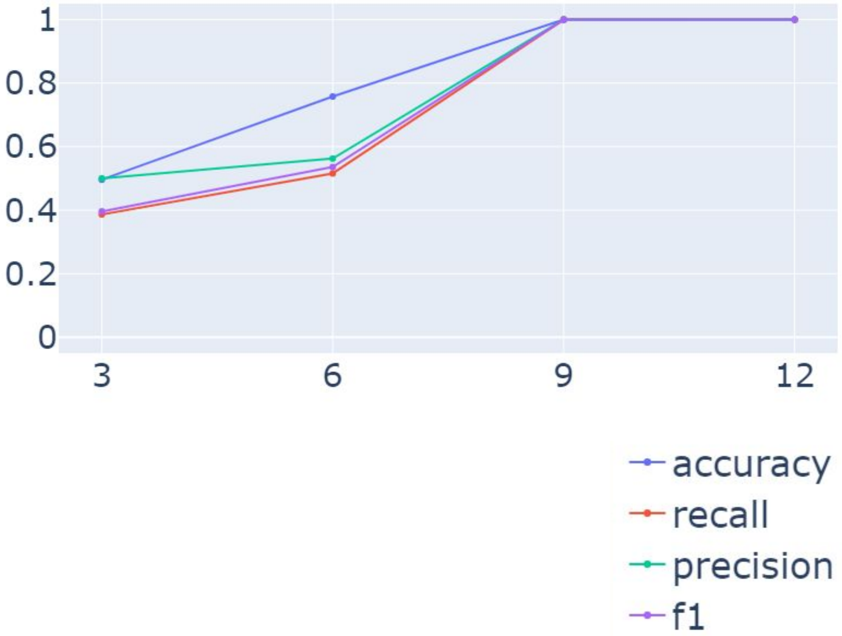


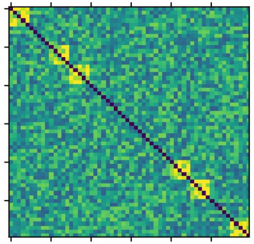


Integration based on **Normal** distribution

metrics depending on *n*

metrics depending on *nar* for n > 1

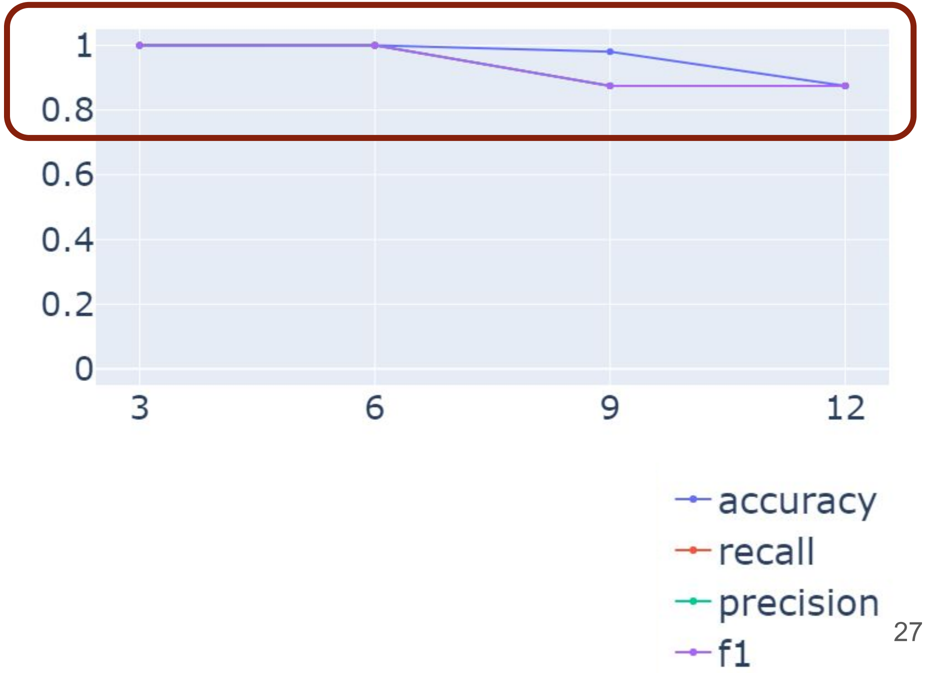


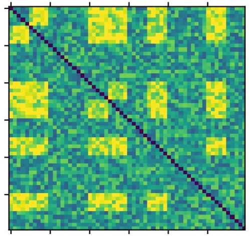


Segregation based on **Uniform** distribution

metrics depending on *n*

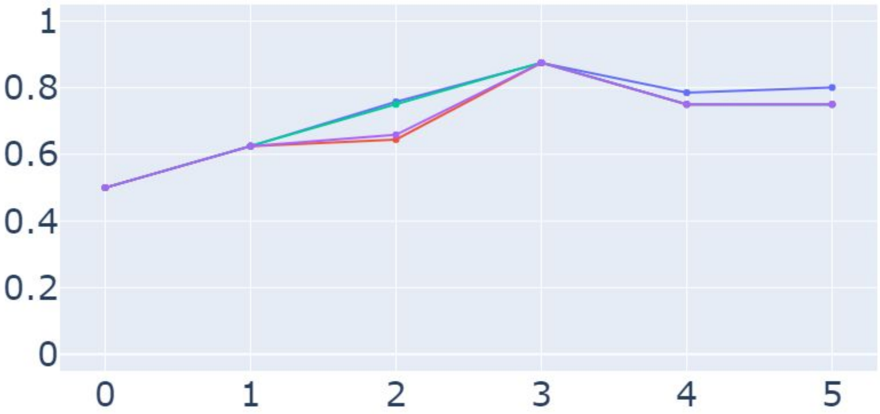
metrics depending on *nar* for n > 1



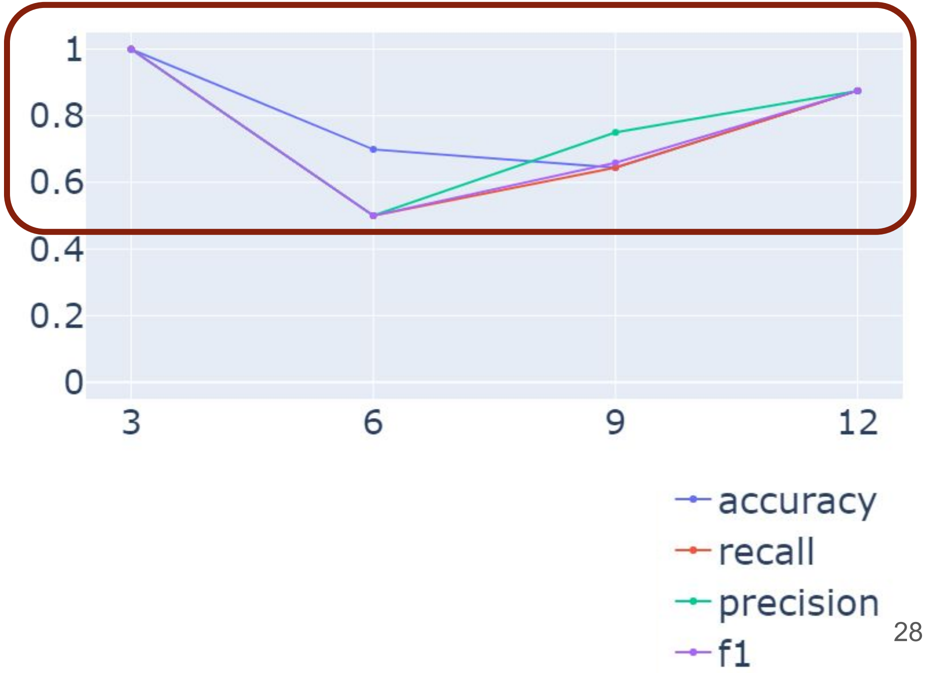


Integration based on **Uniform** distribution

metrics depending on *n*



metrics depending on *nar* for n > 1



📝Conclusion

Normal distribution: it is easier to predict integration rather than segregation.

Segregation: It is easier to predict it based on uniform than on normal distribution.

In most cases performance improves as *n* increases.

Distribution type plays a role in the GNN’s performance.

📚Further work

- Creating not fully connected graphs
- Use an **interpreter**

Acknowledgements

- **Vladimir Maksimenko:** supervision, support, and knowledge sharing
- **Creators of *bctpy* library**