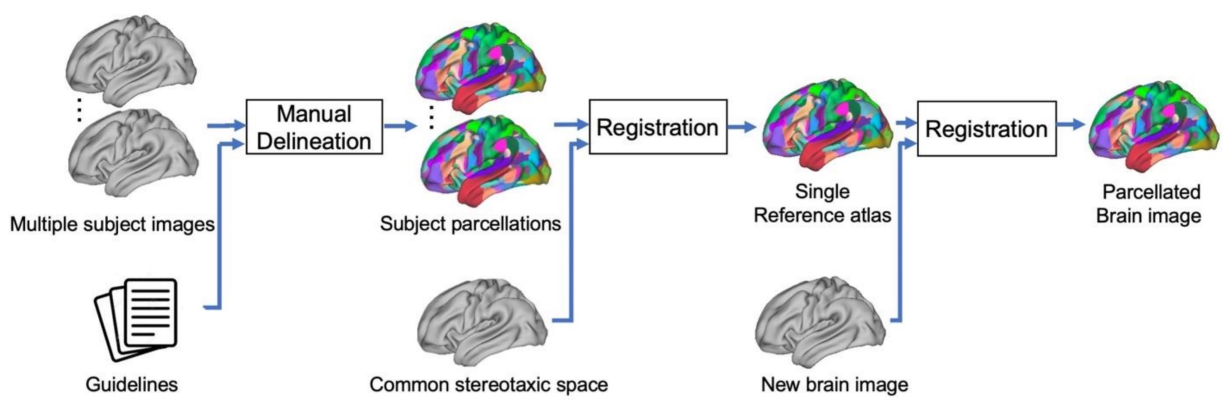
DGL Kaggle Competition 2025: Generative GNNs for Brain Graph Super-resolution

https://github.com/basiralab/DGL/tree/main/Project

Motivation: Why super-resolve graphs?

MRI image processing (Moghimi et al., 2021)



☹ Time-consuming image processing, denoising, registration, and segmentation pipelines.

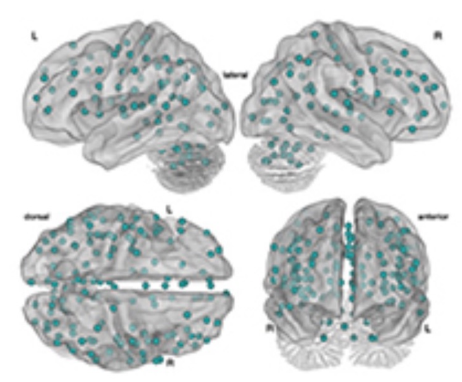
☹ The finer the parcellation (the more nodes we create), the longer the MRI processing time.

☺ Generate directly the brain graph at higher resolutions (with more nodes) while circumventing the need for heavy image data processing.

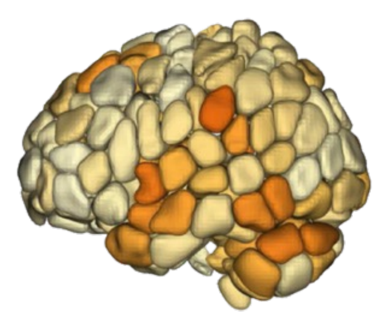
Low-resolution (LR) and high-resolution (HR) brain connectivity data preparation

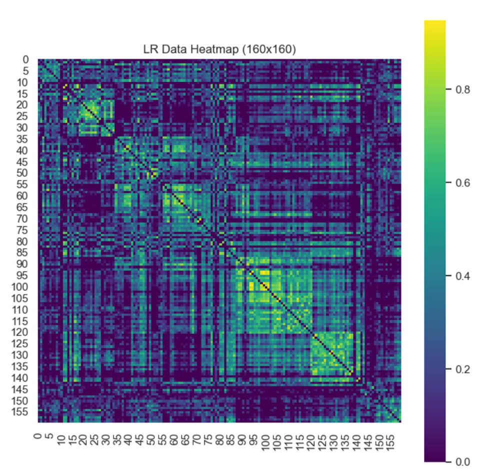
LR and HR parcellation atlases (SLIM dataset, 2017)

Dosenbach 2010 atlas (160 regions/nodes)

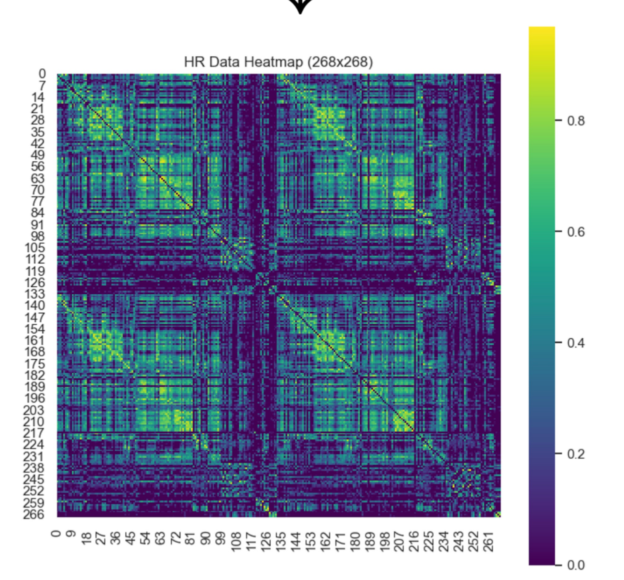


Shen 2009 atlas (268 regions/nodes)



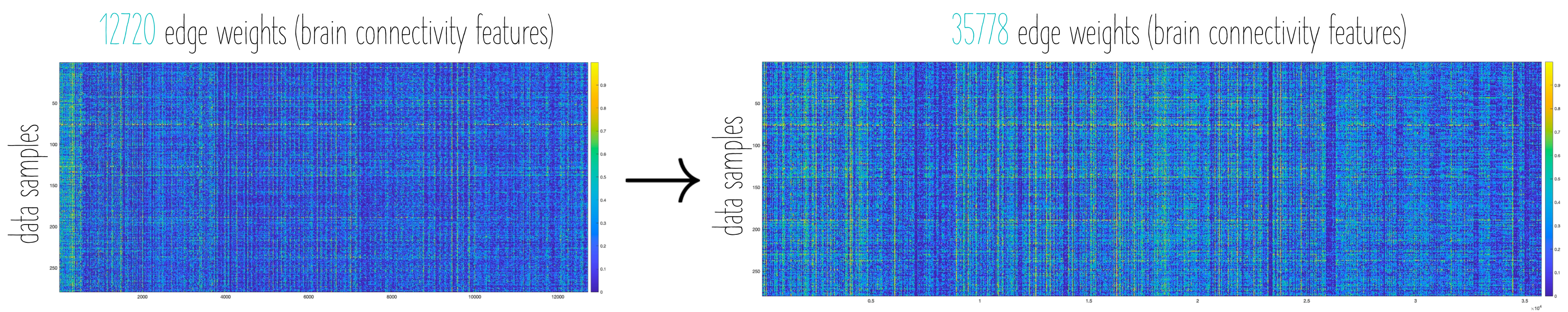


LR (160 nodes) brain graph matrix



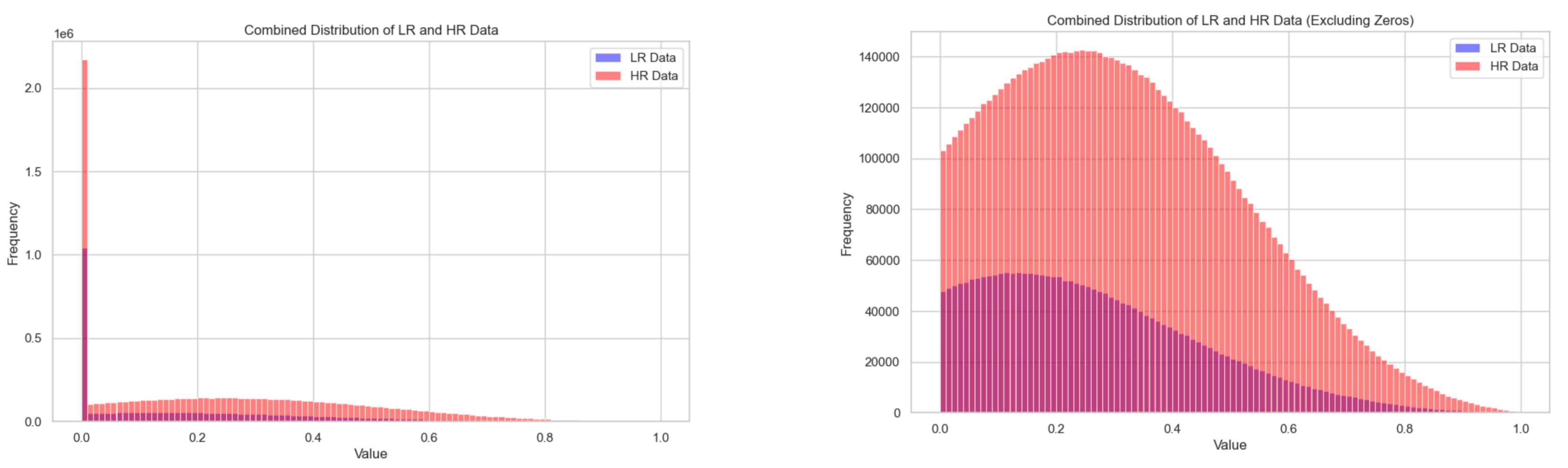
HR (268 nodes) brain graph matrix

Matrix vectorization



Vectorized LR brain graph data

Vectorized HR brain graph data



Data splits for the Kaggle competition

- Training set 167
- Public Test set 56
- Private Test set 56

Train set

167 subjects — $\mathbf{A}^{LR}_{tr}$ — edge weights

167 subjects — $\mathbf{A}^{HR}_{tr}$ — edge weights

Test set

112 subjects — $\mathbf{A}^{LR}_{ts}$

112 subjects — 56 public test / 56 private test

You can see your rank during the competition.

The final rank will be revealed after the submission deadline based on the private test performance.

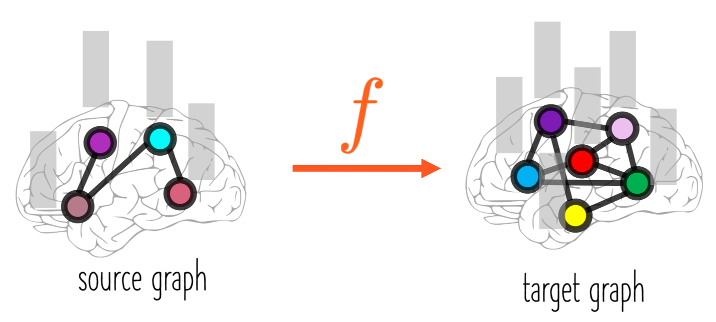
Note: You will perform 3-fold cross-validation on the provided train (LR + HR) data to select and compare different generative models you will use and design.

Submission file format using serialization

The serialization or melting of matrix simply flattens it into a long vector by stacking the rows one by one.

Inductive generative model

1. Define the node features for LR and HR matrices in a consistent manner. The feature size can vary across LR and HR graphs.
2. Anti-vectorize each brain connectivity sample into a full symmetric weighted adjacency matrix. It is an option you can consider.



GNN mapping function

$$f(A, X) = (\tilde{A}', \tilde{X}')$$

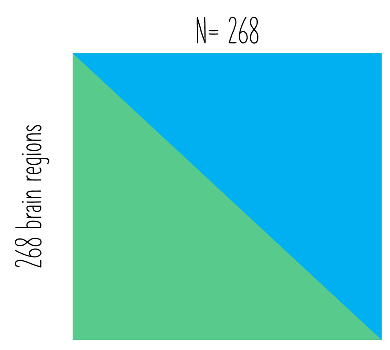
GNN loss function

$$\mathcal{L}_1 = dist(A', \tilde{A}')$$

How to initialize $\boldsymbol{X} = \boldsymbol{H_0}$?

1. Randomly.
2. Using the topological attributes of a given node.
3. Using an identity vector of [1 1 … 1].

Antivectorization

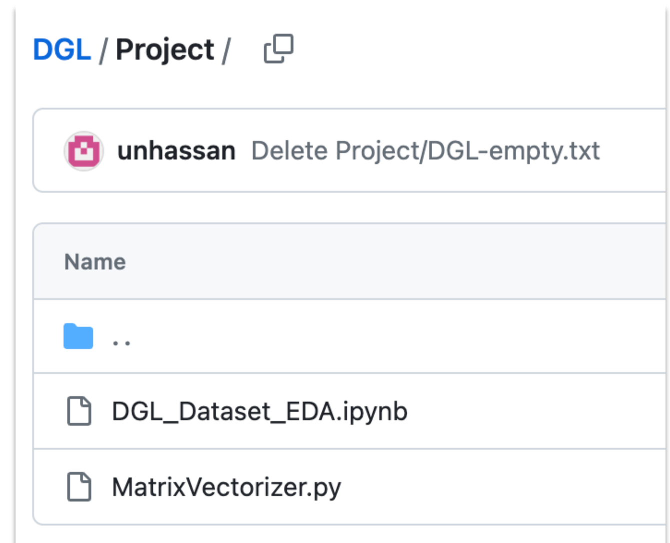


Brain graph adjacency matrix

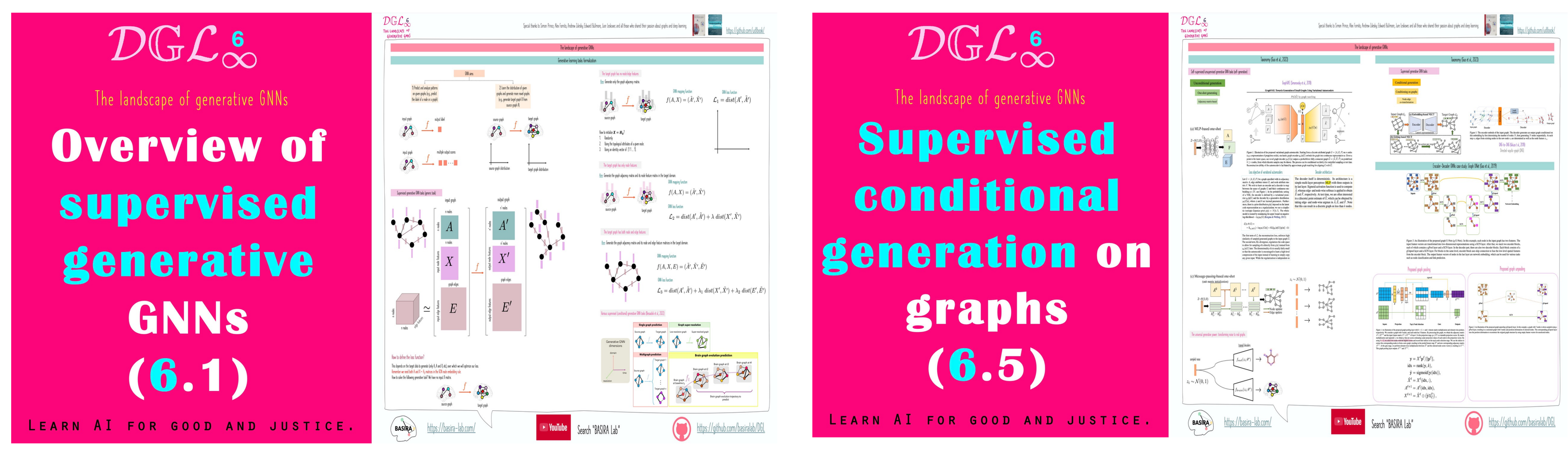
Vectorize vertically the off-diagonal upper triangular part of this matrix (exclude diagonal)

Feature vector of size

N * (N-1)/2 = (268* 267)/2 = 35778



References: DGL lecture 6.1 to 6.7



1. MRI image processing: Moghimi, Pantea, et al. "A Review on MR Based Human Brain Parcellation Methods." arXiv preprint arXiv:2107.03475 (2021).
2. Brain parcellation atlases: https://www.lead-dbs.org/helpsupport/knowledge-base/atlasesresources/cortical-atlas-parcellations-mni-space/
3. Shen 2009 Parcellation: Shen, X., Tokoglu, F., Papademetris, X., Constable, R.T.. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification, Neuroimage, 82, 403-15, 2013.
4. SLIM dataset: Liu, W., et al.: Longitudinal test-retest neuroimaging data from healthy young adults in Southwest China. Sci. Data 4, 170017 (2017)
5. Dosenbach 2010 Parcellation: Dosenbach, N.U., et al.: Prediction of individual brain maturity using fMRI. Science 329, 1358-1361 (2010)
6. Lecture 6 videos: https://www.youtube.com/watch?v=TLiHaXinKIA&list=PLug43ldmRSo14Y_vt7S6vanPGh-JpHR7T&index=24&ab_channel=BASIRALab