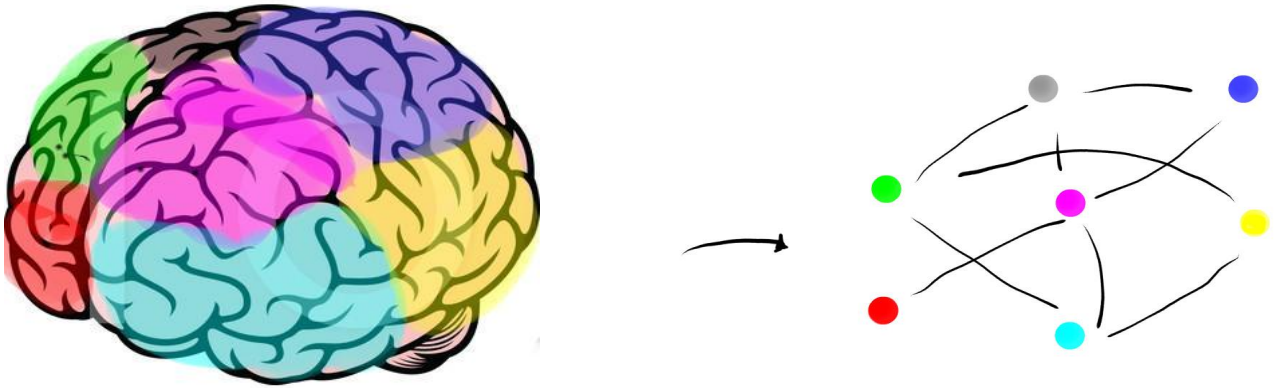
RUSSIAN ACADEMY OF SCIENCES
INSTITUTE FOR INFORMATION TRANSMISSION PROBLEMS
(Kharkevich Institute)

HIGHER SCHOOL OF ECONOMICS
NATIONAL RESEARCH UNIVERSITY

Classification of normal and pathological brain networks based on similarity of graph partitions

Anvar Kurmukov, Yulia Dodonova, Leonid Zhukov

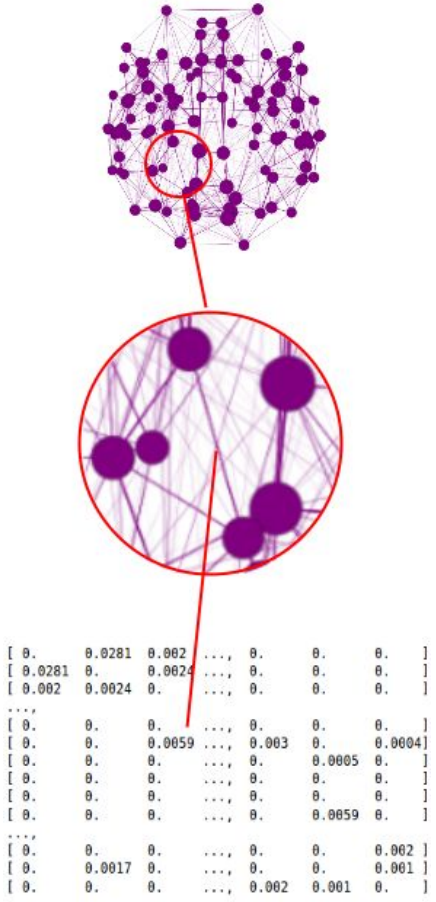
What is a connectome? (connectome = brain network)



At a macroscale, connectome is a graph in which nodes correspond to different brain regions, and edges are the neural connections between these regions

Connectomes: properties

- connectomes are relatively **small** graphs, usually with at most few hundreds of nodes
- the graphs are **undirected**, i.e. the adjacency matrices are symmetric
- edges are **weighted**
- graphs are **connected**
- **each node is uniquely labeled (according to the brain region), and the set of labels is the same across connectomes**
- nodes are localized in **3D space**



```
[ 0.     0.0281  0.002  ...,  0.     0.     0.    ]
[ 0.0281 0.      0.0024 ...,  0.     0.     0.    ]
[ 0.002  0.0024  0.     ...,  0.     0.     0.    ]
...,
[ 0.     0.      0.     ...,  0.     0.     0.    ]
[ 0.     0.      0.0059 ...,  0.003  0.     0.0004]
[ 0.     0.      0.     ...,  0.     0.0005 0.    ]
[ 0.     0.      0.     ...,  0.     0.     0.    ]
[ 0.     0.      0.     ...,  0.     0.     0.    ]
[ 0.     0.      0.     ...,  0.     0.0059 0.    ]
...,
[ 0.     0.      0.     ...,  0.     0.     0.002 ]
[ 0.     0.0017  0.     ...,  0.     0.     0.001 ]
[ 0.     0.      0.     ...,  0.002  0.001  0.    ]
```

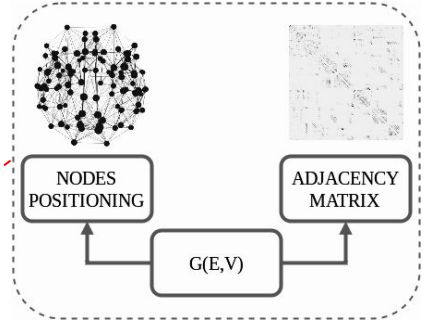
Goal

Given a set of undirected, weighted, connected graphs $X = \{G_1, \ldots G_k\}$, each graph represented by its adjacency matrix $\{A_1, \ldots A_k\}$, we want to predict phenotype (target variable) associated with the graph.

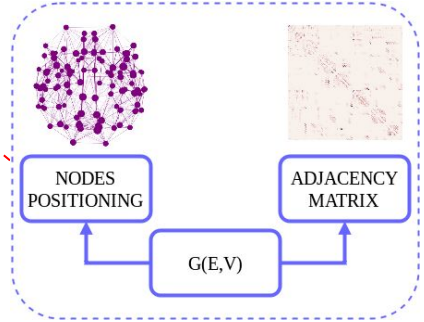
Predict phenotype (e.g., normal or pathological development) of the new unseen brain based on the given examples

We consider a binary classification task:
for each graph target variable is either *0* or *1*

Example of Phenotype I

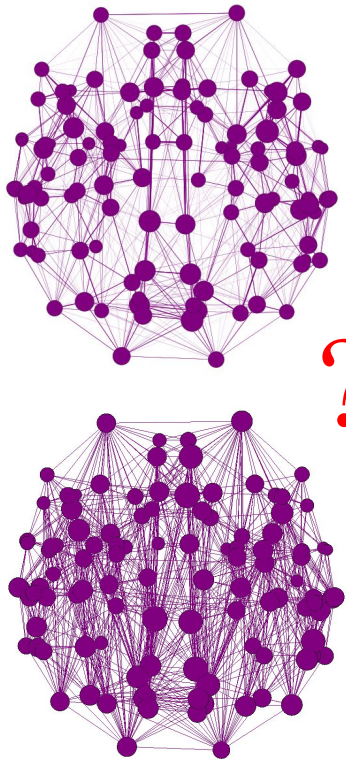


Example of Phenotype II



How to classify graphs?

Problem: Methods of supervised learning usually work with vectors, not graphs



- **Graph embedding methods**

  Describe a network via a vector, *nothing about this approach today*

- **Kernel classifiers**

  Define a positive semi-definite function (**kernel**) on graphs and feed the resulting Gram matrix to the SVM (support vector machines)

  If we introduce a distance ***d(G,G')*** between the two graphs, a kernel can be produced by:

$$K(G, G') = e^{-\alpha d(G,G')}$$

How to compute a distance between two connectomes?

Idea

Connectomes obtained from normal and pathological brains might differ in how brain regions cluster into communities

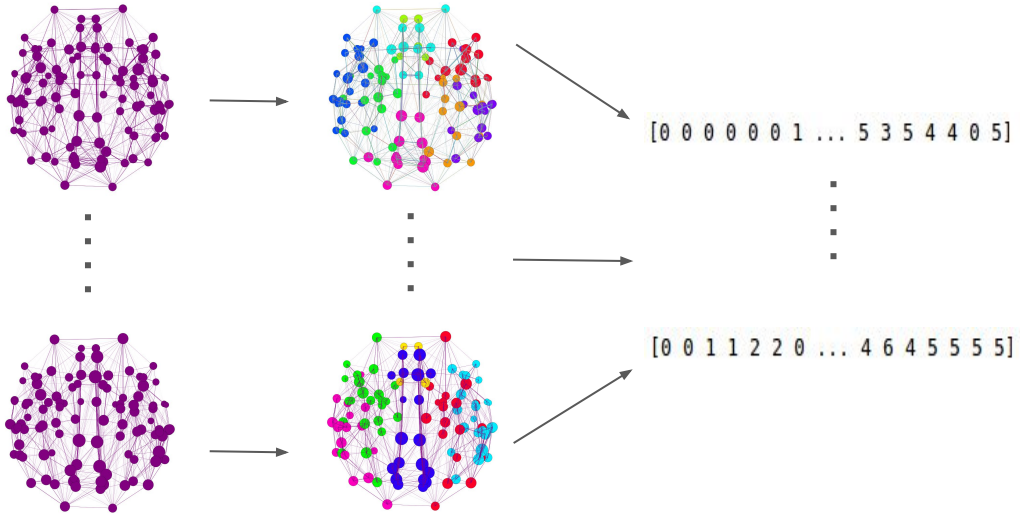
For each brain network, find its **best partition** into clusters

We expect these partitions to be **similar** between brain networks that belong to the **same class** (normal or pathological) and **differ across classes** (between subjects with and without brain disease)

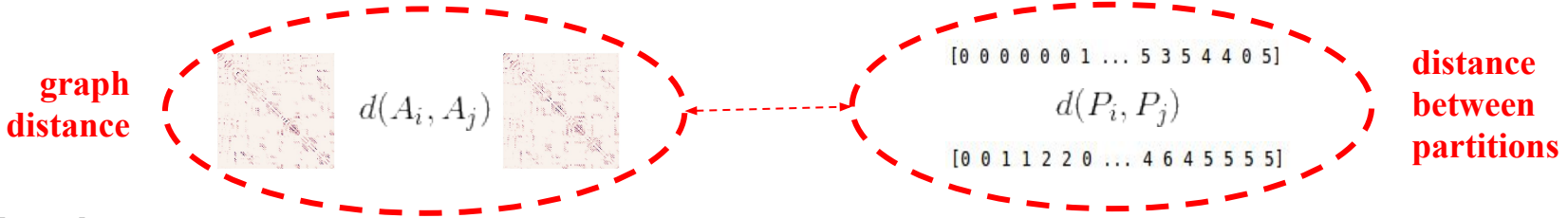
We measure a distance between graphs as a **distance between their partitions**

Similarity of graph partitions

For each graph, we obtain its best partition P which is a vector of length n, where n is the number of nodes.
i-th value in P represents community label of an i-th node.

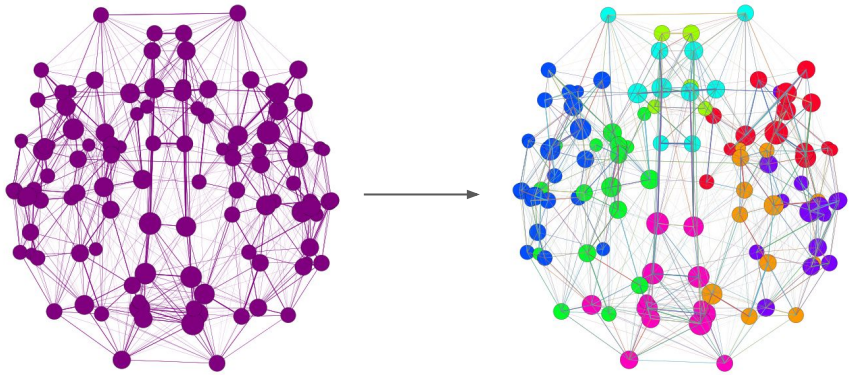


Given a set of graphs $X = \{G_1, \ldots G_k\}$, we obtain partitions $\{P_1, \ldots P_k\}$. Now we want to compare graphs based on similarity in their partitions into communities.



Methods for graph partitioning

- Approximate
  - **Newman eigenvector**
  - **Louvain**
  - **Greedy modularity optimization**
    - *Very fast*
    - *Suboptimal*
- Exact modularity optimization
  - *Computationally hard*
  - *Global modularity optimum*



All algorithms optimize modularity Q which is given by the formula:

$$Q = \frac{1}{2m} \sum_{i,j} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(i,j)$$

1. Newman, M. E. J. (2006) Finding community structure in networks using the eigenvectors of matrices, Phys. Rev. E, 74, 036104.
2. Blondel, V.D., Guillaume, J.-L., Lambiotte, R., Lefebvre, R. (2008) Fast unfolding of communities in large networks, Journal of Statistical Mechanics: Theory and Experiment, 10, P10008.
3. Clauset, A., Newman, M. E. J., Moore, C. (2004) Finding community structure in very large networks. Phys Rev E, 70, 066111 .

Similarity between partitions

- **Adjusted Rand Index**

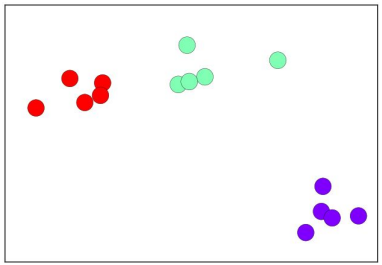
$ARI(P_1, P_1) = 1.0$
$ARI(P_1, P_2) = 1.0$
$ARI(P_1, P_3) = 0.479$
$ARI(P_1, P_4) = 0.042$

Both ARI and AMI are indifferent to cluster relabeling

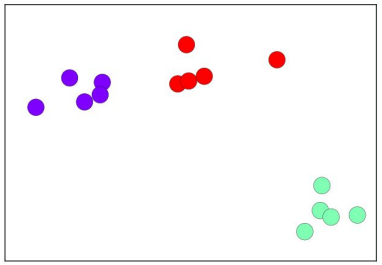
- **Adjusted Mutual Information**

$AMI(P_1, P_1) = 1.0$
$AMI(P_1, P_2) = 1.0$
$AMI(P_1, P_3) = 0.529$
$AMI(P_1, P_4) = 0.049$

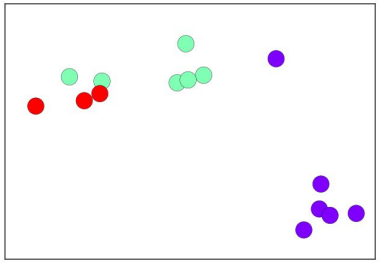
Partition 1 : [0 0 0 0 0 1 1 1 1 1 2 2 2 2 2]



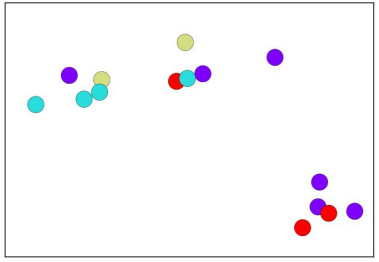
Partition 2 : [1 1 1 1 1 2 2 2 2 2 0 0 0 0 0]



Partition 3 : [0 0 0 0 0 0 1 1 1 1 1 1 2 2 2]



Partition 4 : [0 0 0 3 3 0 2 0 3 1 2 0 1 1 1]

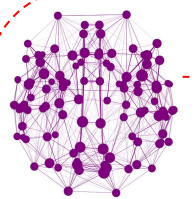


Both ARI and AMI take the value 1 when two partitions are identical and values close to 0 for random labeling

Take (1-ARI) and (1-AMI) to obtain distances

Vinh, N. X., Epps, J., & Bailey, J. (2010). Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance. *Journal of Machine Learning Research*, *11*(Oct), 2837-2854.

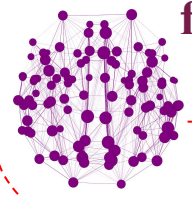
Classification pipeline



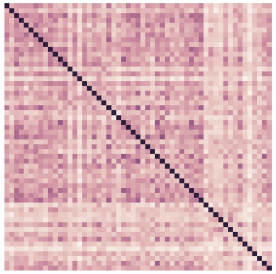
[0 0 0 0 0 0 1 ... 5 3 5 4 4 0 5]

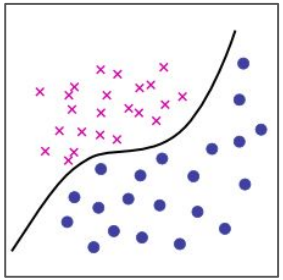
Step 1
obtain best partition for each graph

Step 2
compute pairwise distances between partitions



[0 0 1 1 2 2 0 ... 4 6 4 5 5 5 5]





Step 3
produce a kernel by exponentiating distances and run kernel SVM

Produce baseline classification quality:

Step 1
compute pairwise *L1* (Manhattan) and *L2* (Euclidean) distances between graph adjacency matrices

Data

Phenotypes: Carriers versus non-carriers of the APOE-4 allele associated with the higher risk of Alzheimer's disease.

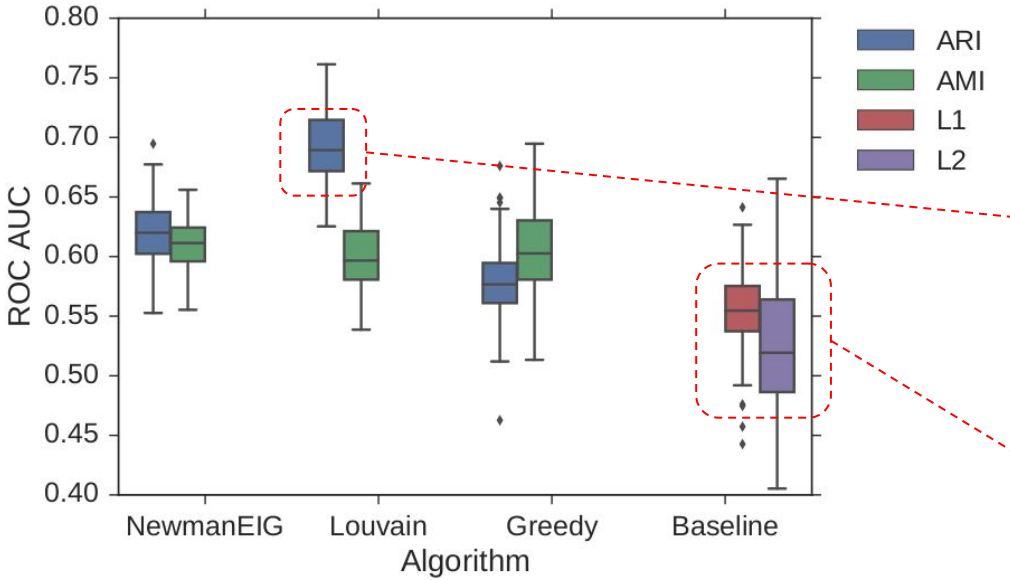
Dataset: Publicly available UCLA APOE-4 dataset (UCLA Multimodal Connectivity Database), includes precomputed DTI-based matrices of structural connectomes. The sample includes

Basics: 30 APOE-4 non carriers, mean age (age standard deviation) is 63.8 (8.3), and 25 APOE-4 carriers, mean age (age standard deviation) is 60.8 (9.7).

Classification pipeline: summary

- Compute graph partitions using **three different algorithms**
  - *Newman eigenvector*
  - *Louvain*
  - *Greedy modularity optimization*
- Compute partition similarities using **two similarity measures**
  - *Adjusted Rand Index*
  - *Adjusted Mutual Information*
- Produce **kernels** from similarity matrices
- Use **SVM** for classification
- Use **10-fold cross-validation** procedure (results averaged over 100 different 10-fold splits)

Results

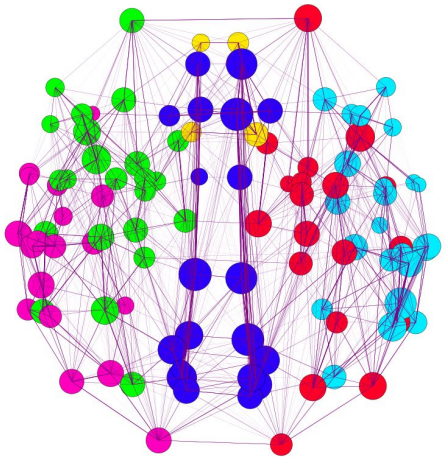


Kernel that uses
Louvain partitioning + ARI similarity

Baseline kernels based on *L1* and *L2* distances between the graph adjacency matrices

Best result is obtained with Louvain partitioning and Adjusted Rand Index. SVM classifier with this kernel clearly **outperforms the baseline** and gives ROC AUC **0.7 ± 0.03** (mean ± std).

Conclusions



- Network science is becoming a popular instrument for neuroscience research: neural connections of a human brain are modeled by a **graph called connectome**
- A task is to **classify** these small undirected graphs
- Idea: if the connectomes come from the same class, their nodes (brain regions) **cluster into communities similarly**
- Hence, measure **distances** between connectomes based on similarity in partitions, construct a kernel based on these distances and use a kernel classifier
- This approach **outperforms** kernels based on simple distances between the adjacency matrices of the respective graphs *(shown today)* and graph embedding methods *(not shown)*

Thank you!

Q?

kurmukovai@gmail.com

Classification of normal and pathological brain networks based on similarity of graph partitions

Anvar Kurmukov, Yulia Dodonova, Leonid Zhukov