Spectral Graph Convolutions for Population-based Disease Prediction

Published in 2017 at MICCAI :
International Conference on Medical Image Computing and Computer-Assisted Interventions

Sarah Parisot, Sofia Ira Ktena, Enzo Ferrante, Matthew Lee, Ricardo Guerrerro Moreno, Ben Glocker, Daniel Rueckert

Manal Akhannouss, Balthazar Neveu, Inès Vati

Paper overview

Disease prediction

Contribution: Graph Convolution Networks [*]

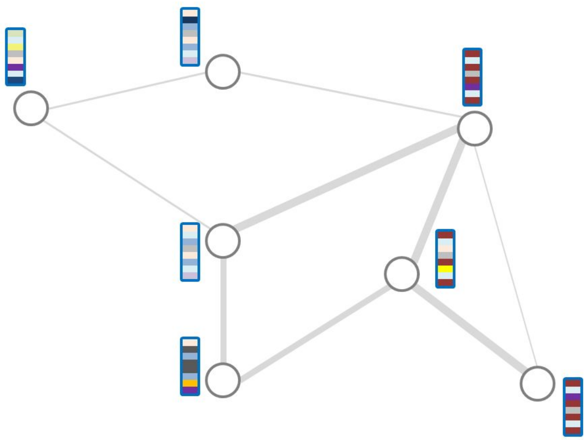
Datasets : ABIDE → Autism | ADNI → Alzheimer

Population graph

Nodes = Patients

Edges = Similarity between patients

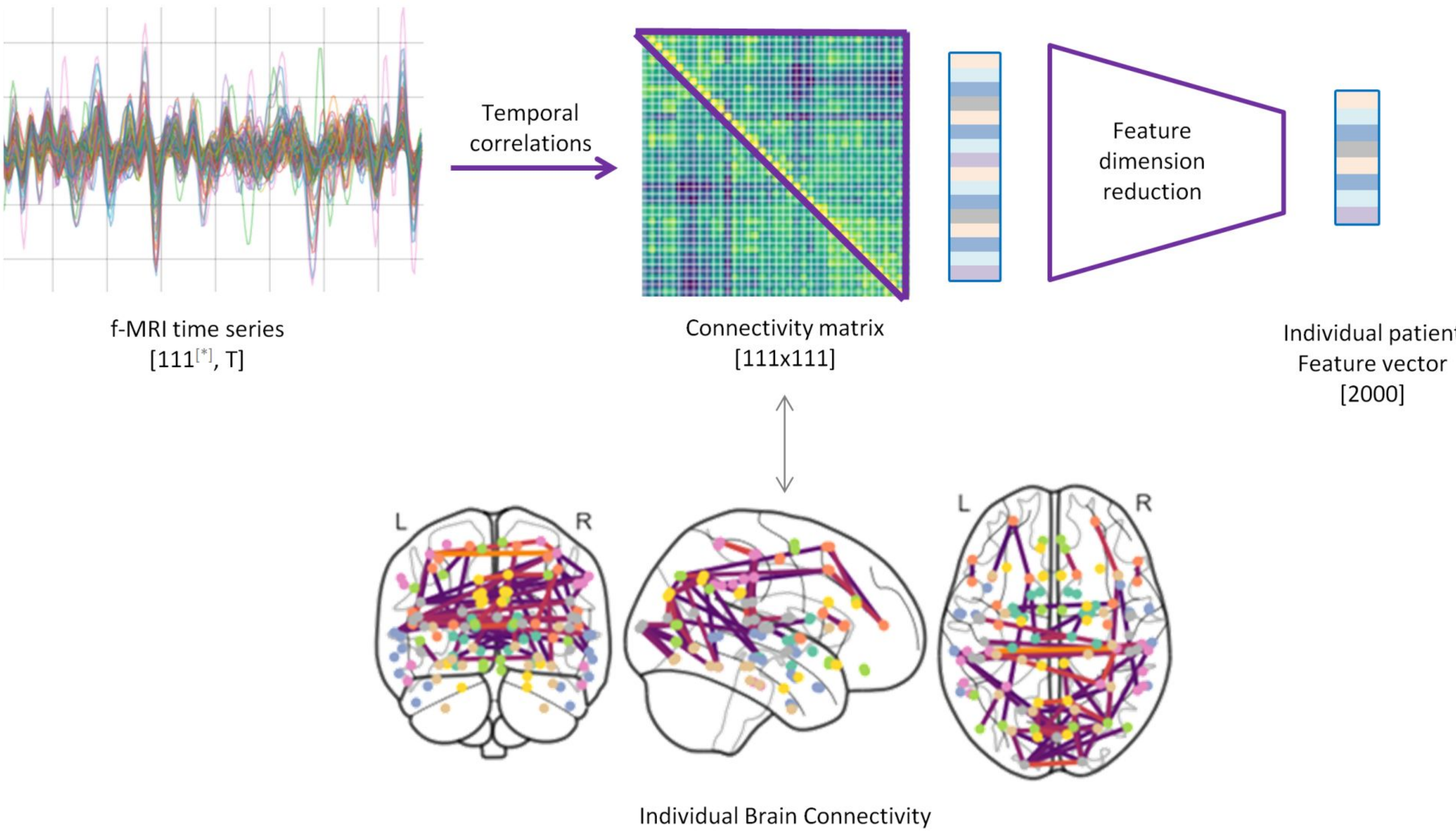
Feature vector



[*] Semi-Supervised Classification with Graph Convolutional Networks - Thomas N. Kipf, Max Welling ICLR **2017**

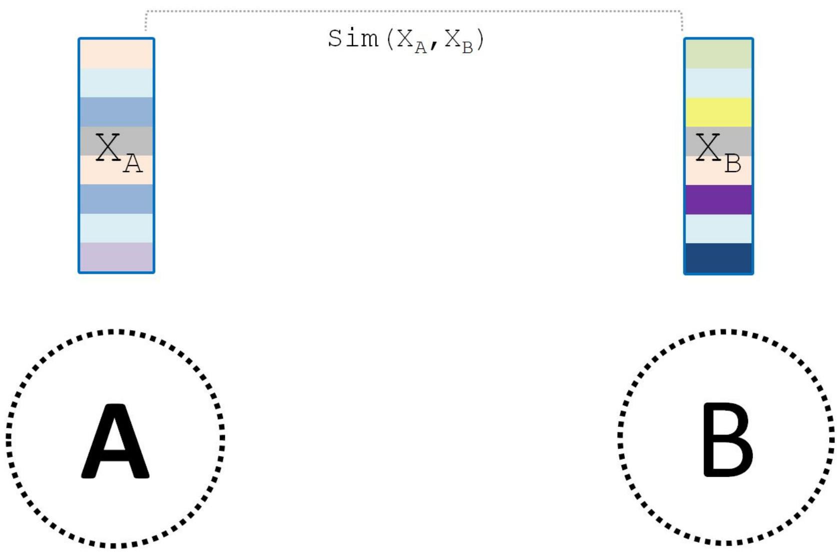
Defining graph nodes

1 patient ↔ 1 node

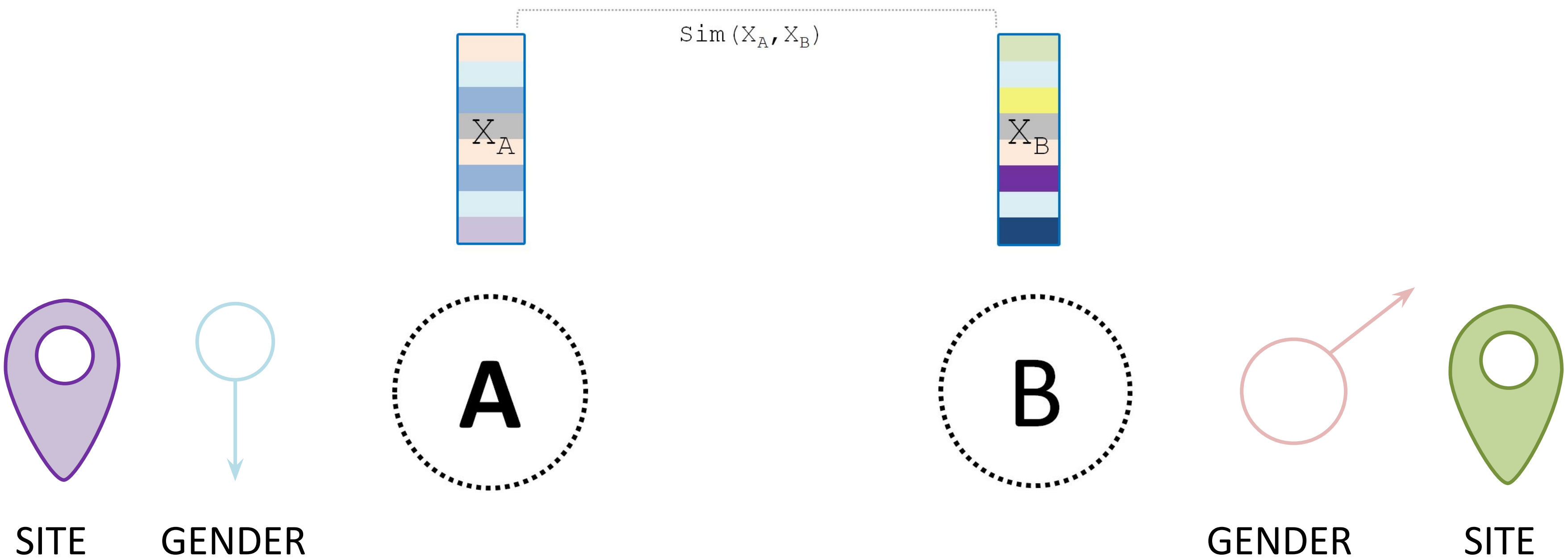


[*] 111: Harvard-Oxford cortical and subcortical structural atlases

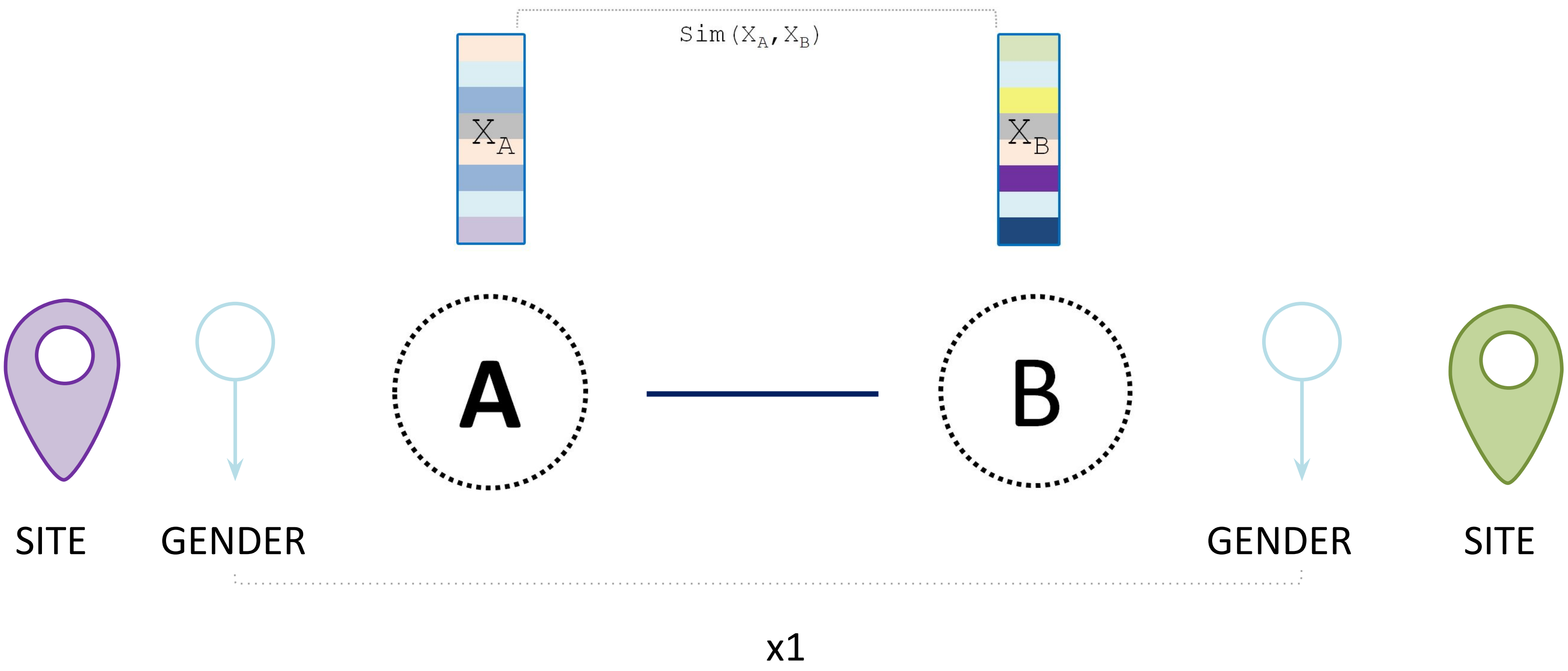
Defining graph edges



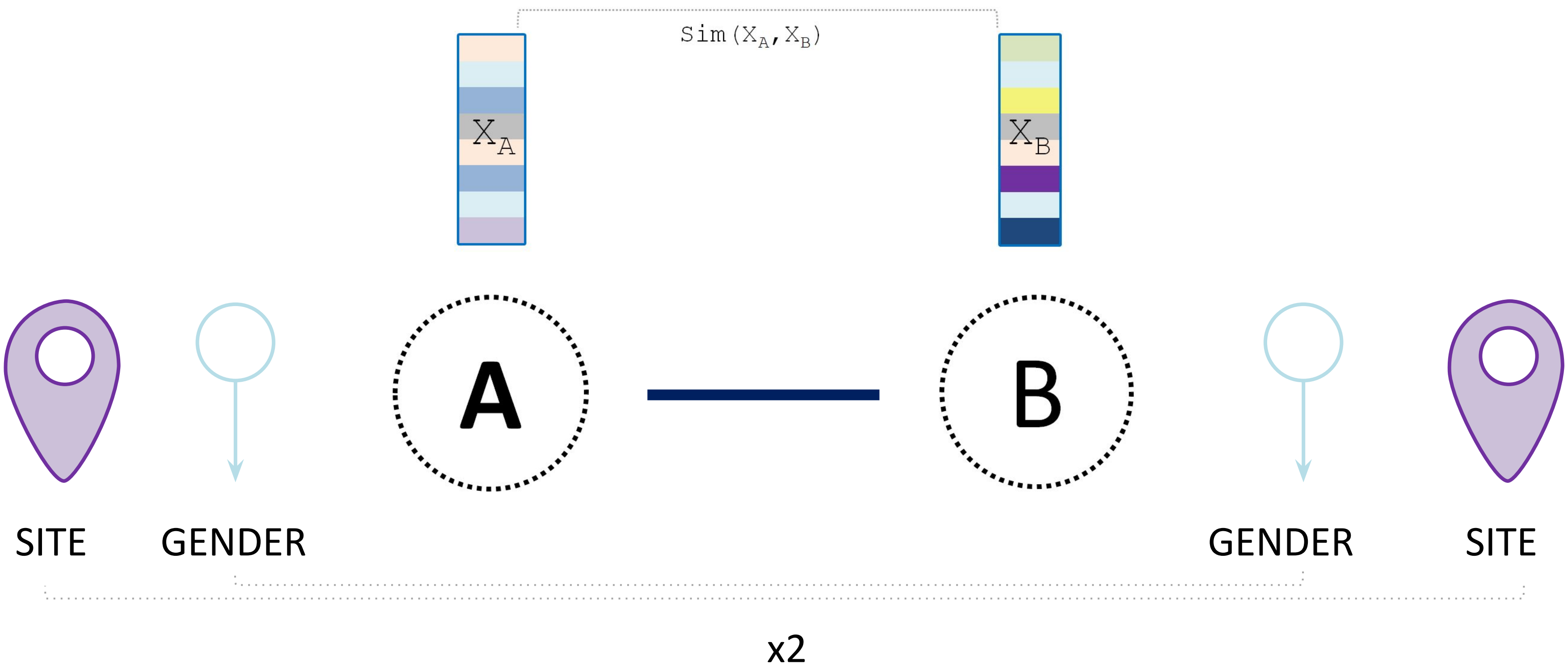
Defining graph edges



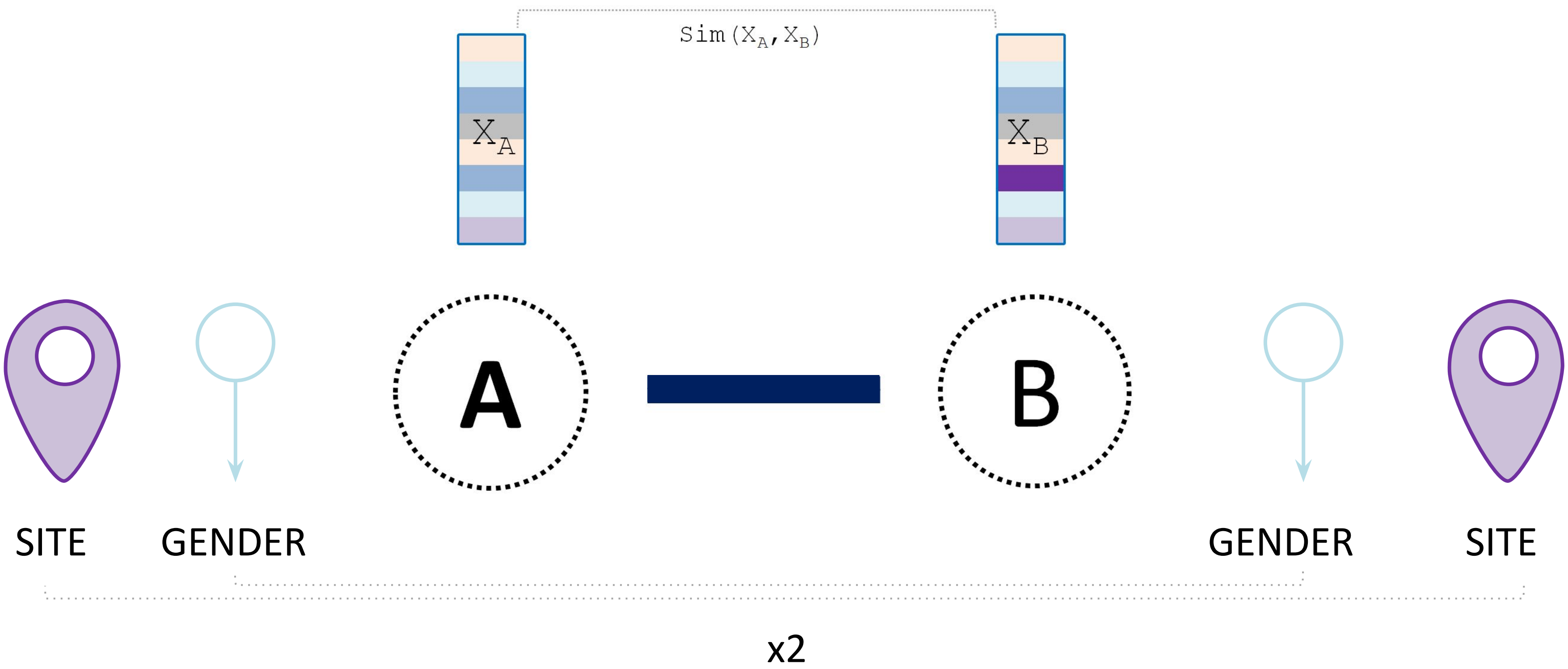
Defining graph edges



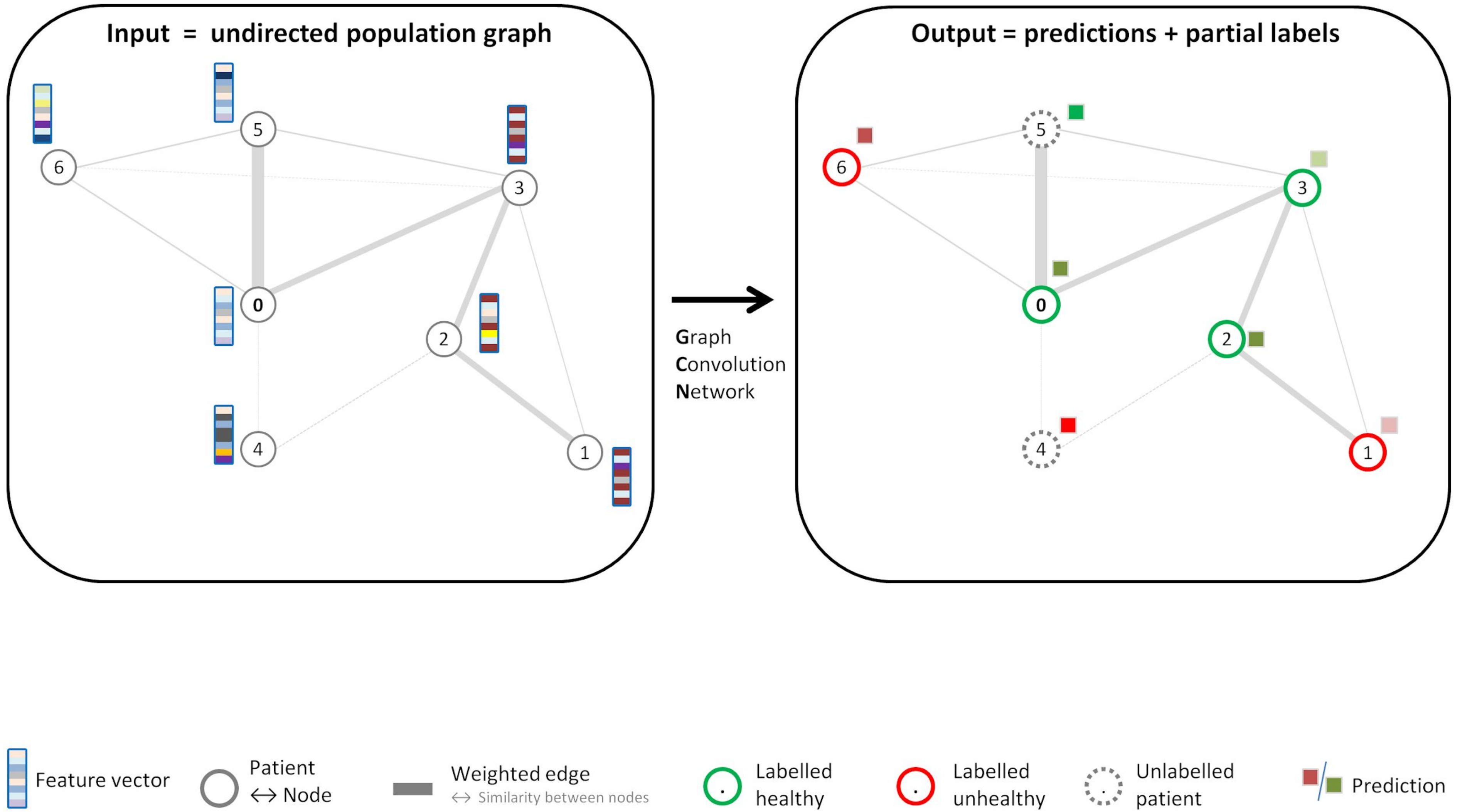
Defining graph edges



Defining graph edges

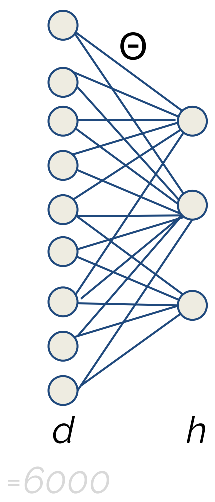


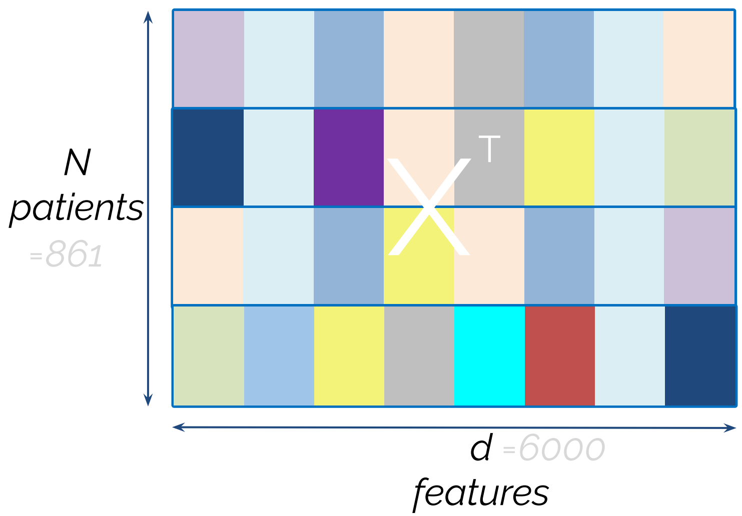
Graph Convolution Network



Graph Convolution **Network**

$$\underbrace{Y}_{h,N} = \underbrace{\Theta}_{h,d} \underbrace{X}_{N,d}{}^{T}$$



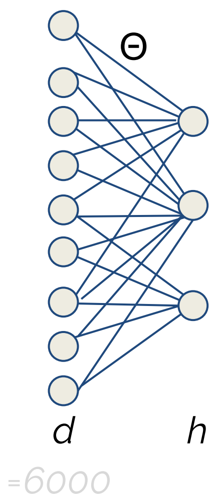


Fully connected *[h,d]*

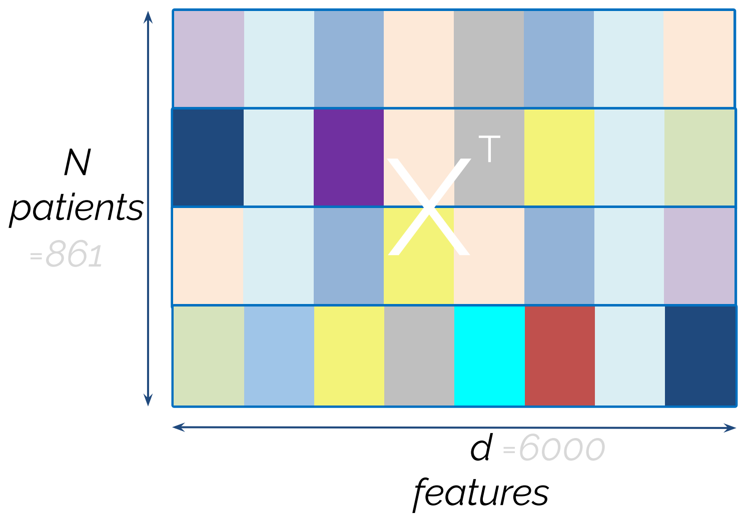
Input features *[N, d]*

Graph Convolution Network

$$\underbrace{Y}_{h,N} = \underbrace{\Theta}_{h,d} \underbrace{X}_{N,d}{}^{T} \underbrace{A}_{N,N}$$

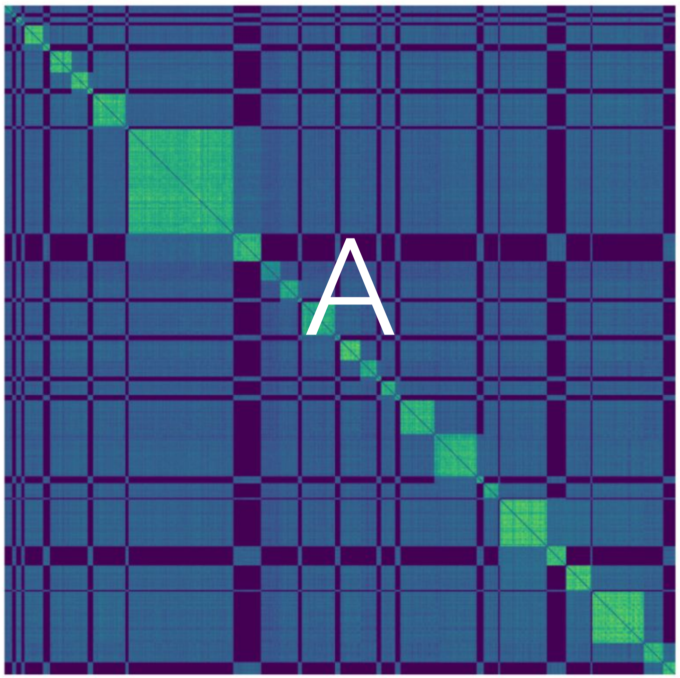


Fully connected *[h,d]*



Input features *[N, d]*

Basic message passing

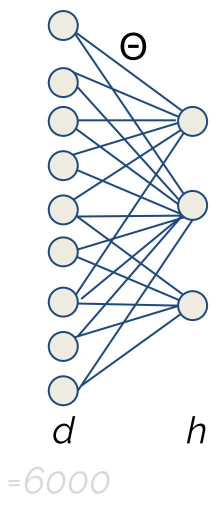


Adjacency matrix *[N, N]*

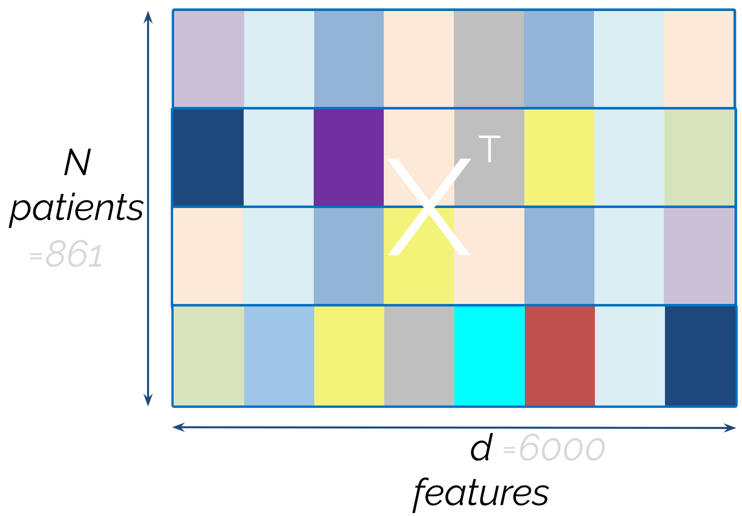
Analogy: Non Local means - denoise similar colors on images

Graph Convolution Network

$$\underbrace{Y}_{h,N} = \underbrace{\Theta}_{h,d} \underbrace{X}_{N,d}{}^{T} \underbrace{A}_{N,N}$$

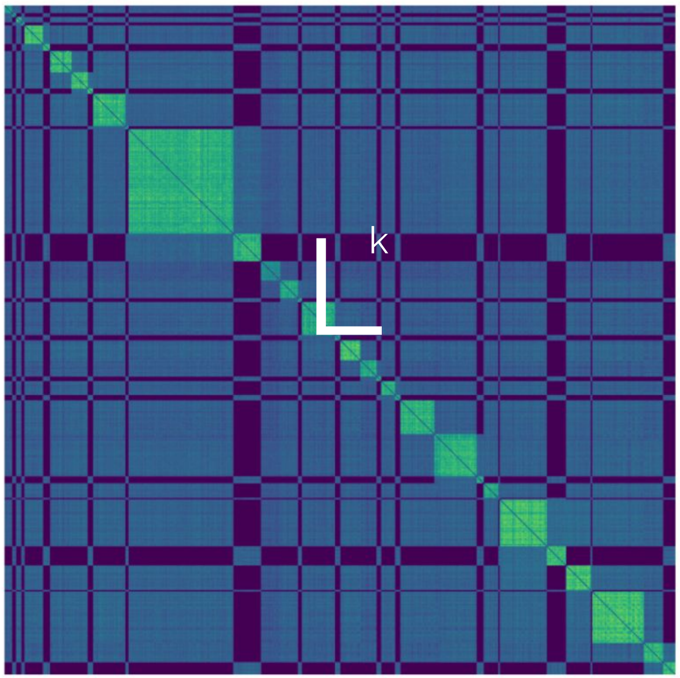


Fully connected *[h,d]*



Input features *[N, d]*

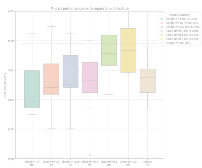
Basic message passing



Power of Laplacian matrix *[N, N]*

Variant in the paper : ChebConv - polynomial of the Laplacian L=I-A ➡ gives access to a larger neighborhood

Experimental setup



Accuracy 68.9% +/- 4% on ABIDE (Autism Disorder Spectrum)
Ability to reproduce results (~69.5%)
→ Compare various models

Cross validation over 10 runs
80% train, 10% validation, 10% test

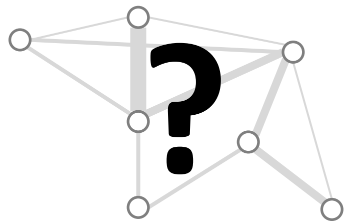
Frozen hyperparameters
Adam, LR 1E-4, Weight decay 0.1, no LR scheduler, 1000 epochs

Reproducible results

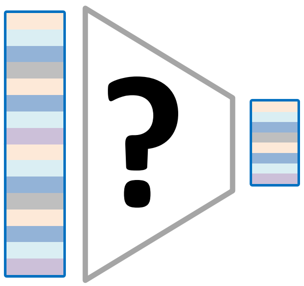
Code from scratch with documentation.

Results

- Is the use of graphs truly relevant?
  Compare
  - Dense Neural Networks
  - Graph Convolutional Networks **+5%**



- Is input dimension reduction relevant?
  No significant differences between
  - Raw input features
  - Recursive Feature Elimination **+0%**
  - Auto Encoders' Latent vector **-3%**

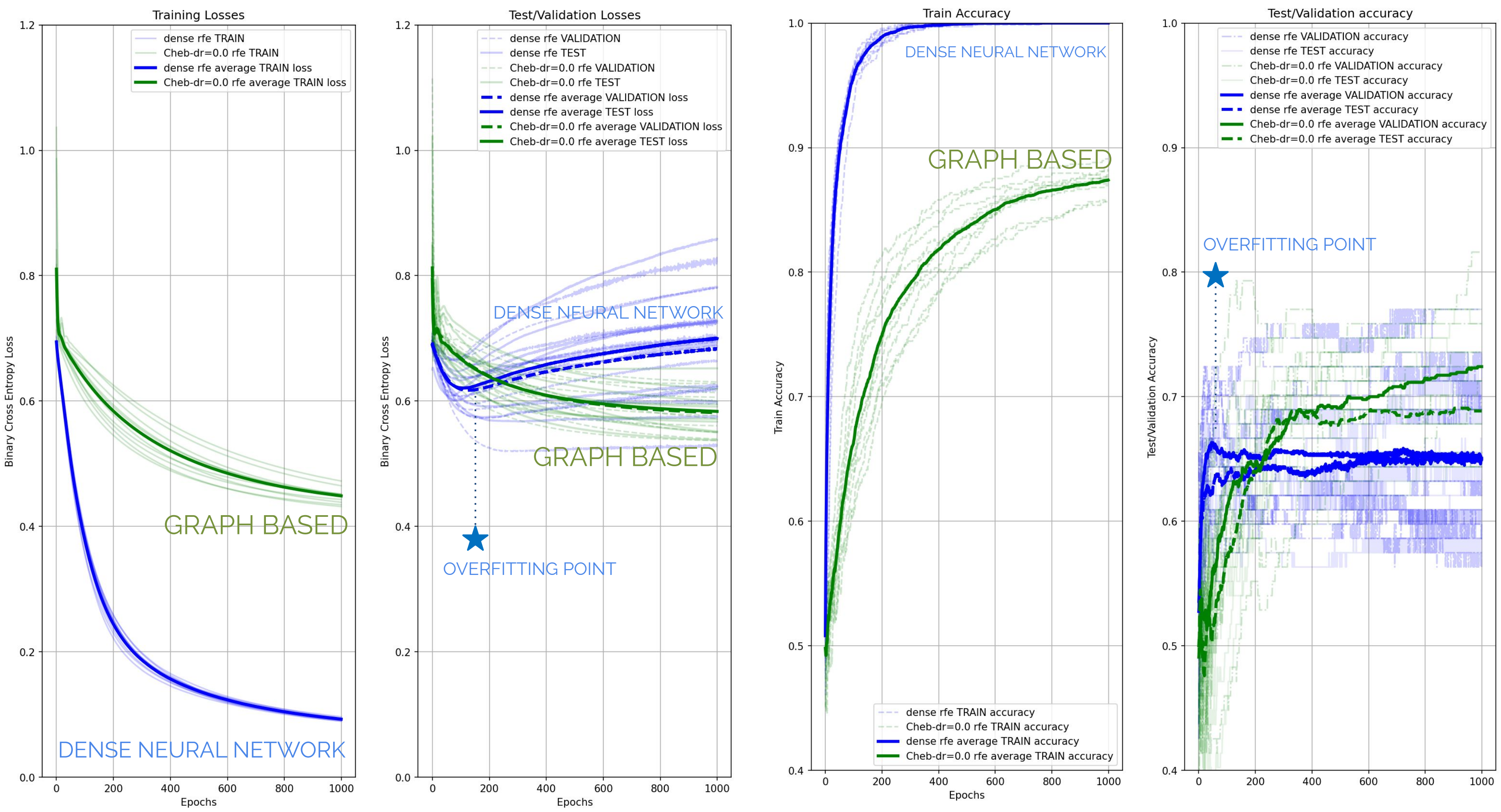


Conclusion

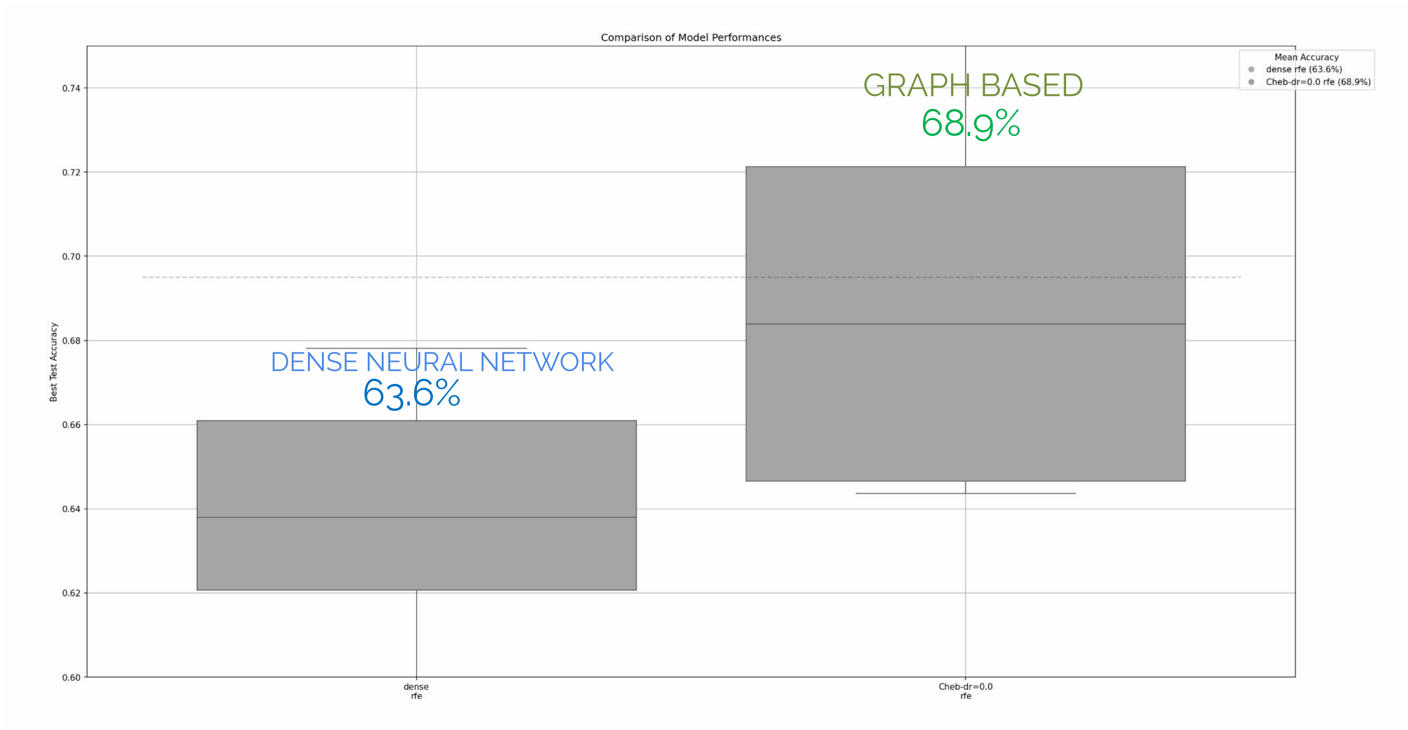
- A good introduction to the difficult domain of medical data analysis
- Curse of dimensionality
- Personal insights
  - Data type : time series , not 4D f-MRI volumes
  - Memory requirements : laptop GPU Nvidia T500 4Gb
  - « Spectral graph convolutions »

Thank you

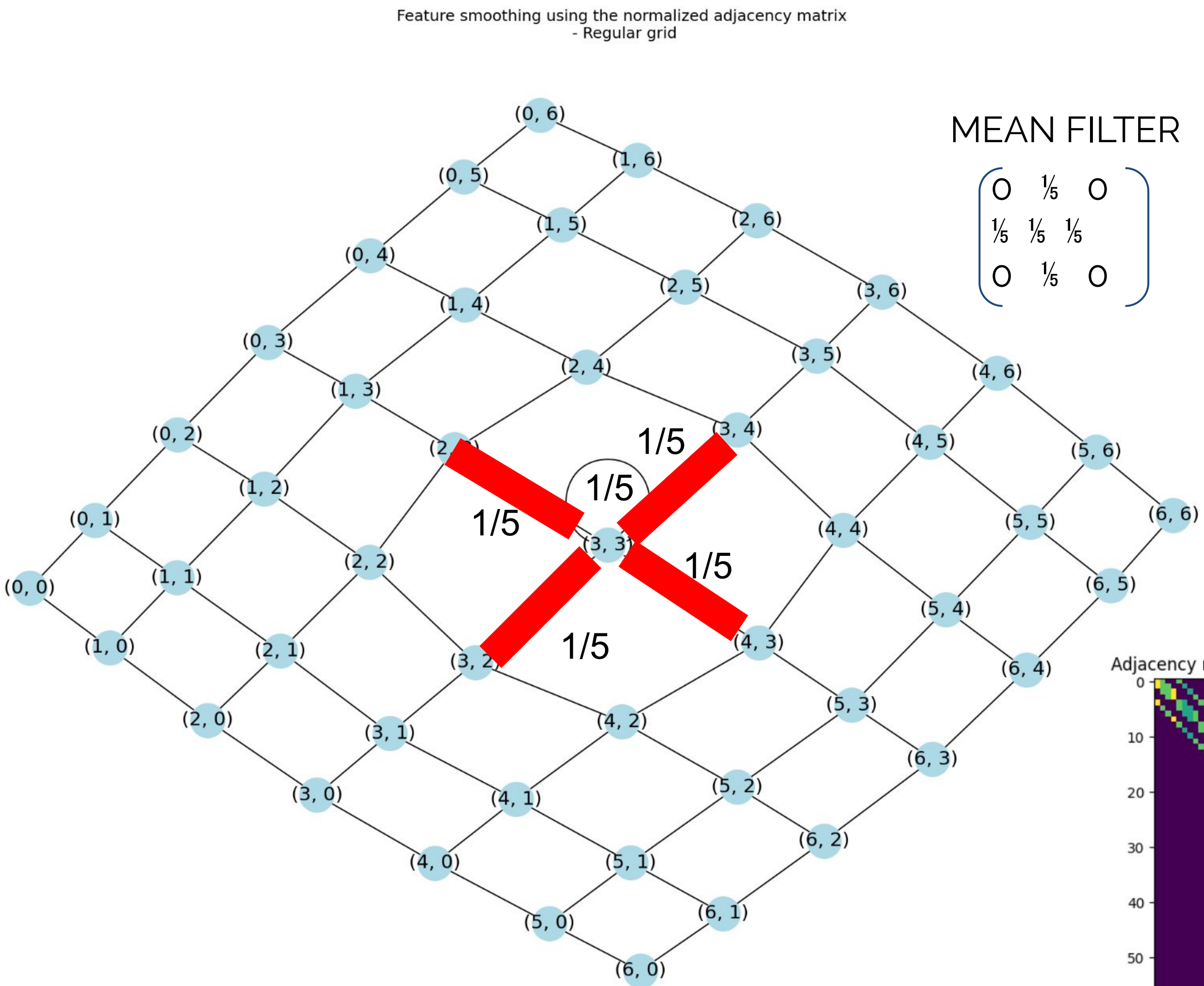
Training curves



Dense Neural Network VS Graph Convolution

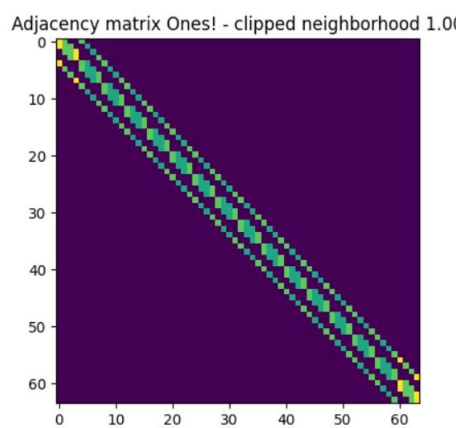


Analogy : Graph convolutions and Image denoising



MEAN FILTER

$$\begin{pmatrix} 0 & \frac{1}{5} & 0 \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ 0 & \frac{1}{5} & 0 \end{pmatrix}$$



Analogy : Graph convolutions and Image denoising

CONSTANT WEIGHTS ON EDGES

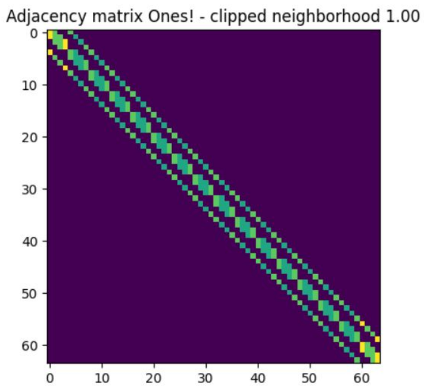
MEAN FILTER

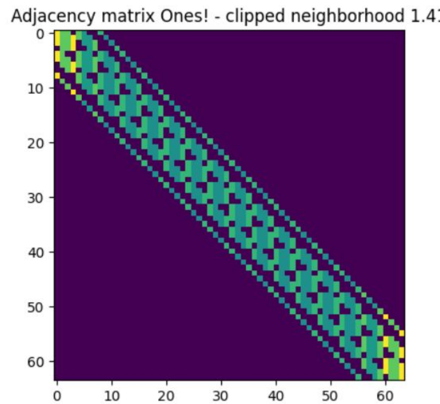
$$\begin{pmatrix} 0 & 1/5 & 0 \\ 1/5 & 1/5 & 1/5 \\ 0 & 1/5 & 0 \end{pmatrix}$$

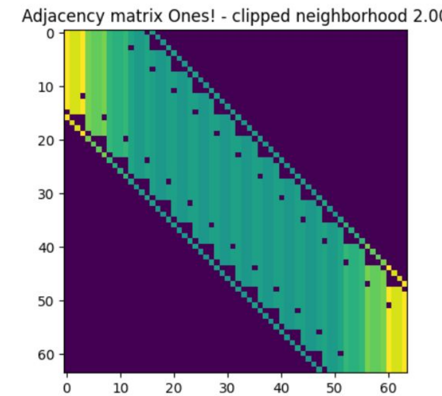
BOX FILTER

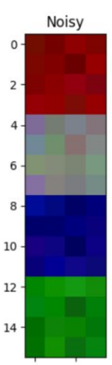
$$1/9 \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

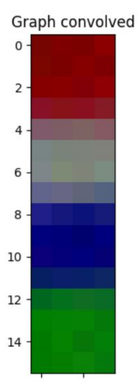
$$1/13 \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

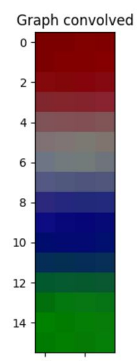


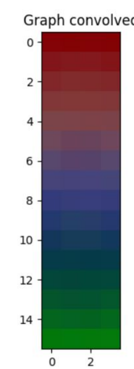








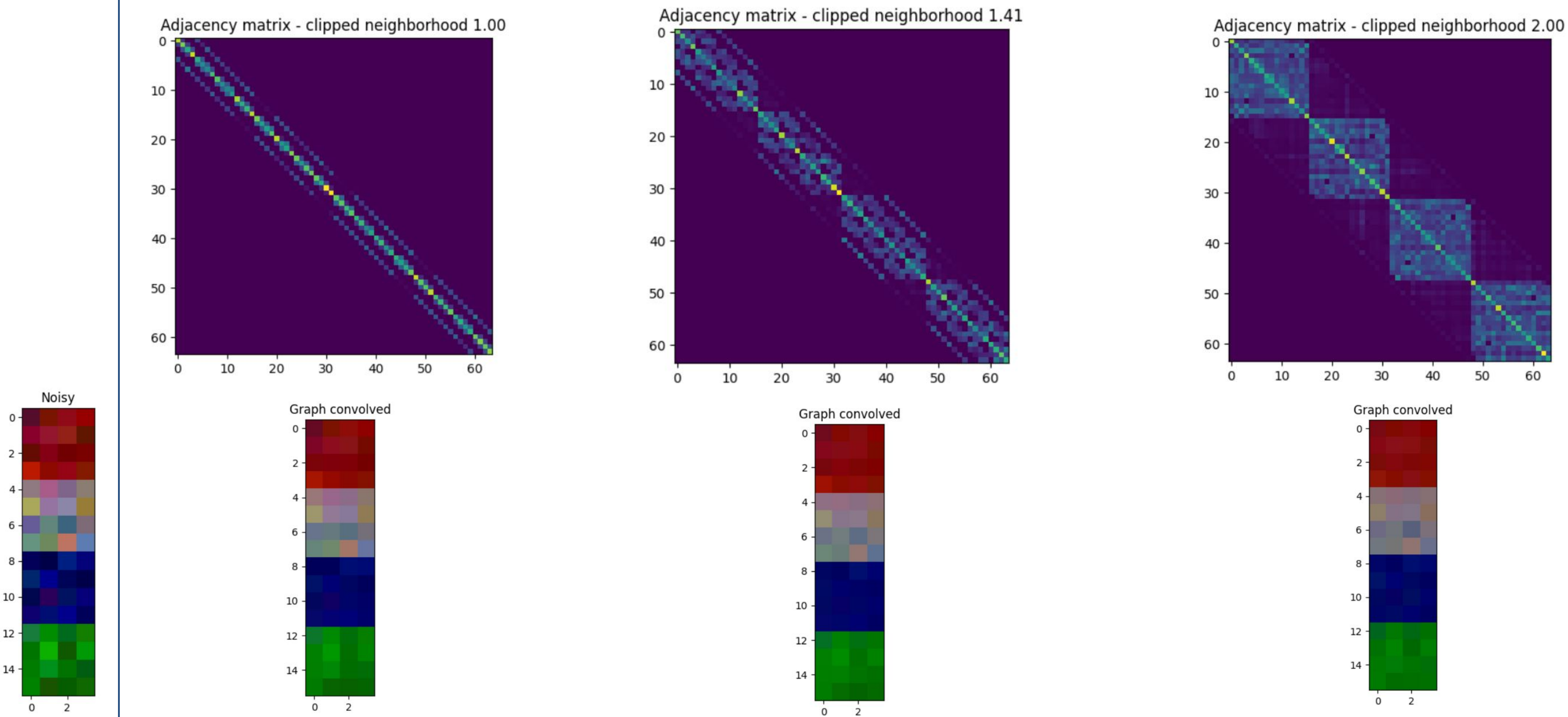




Analogy : Graph convolutions and Image denoising

SIMILARITY WEIGHTS ON EDGES

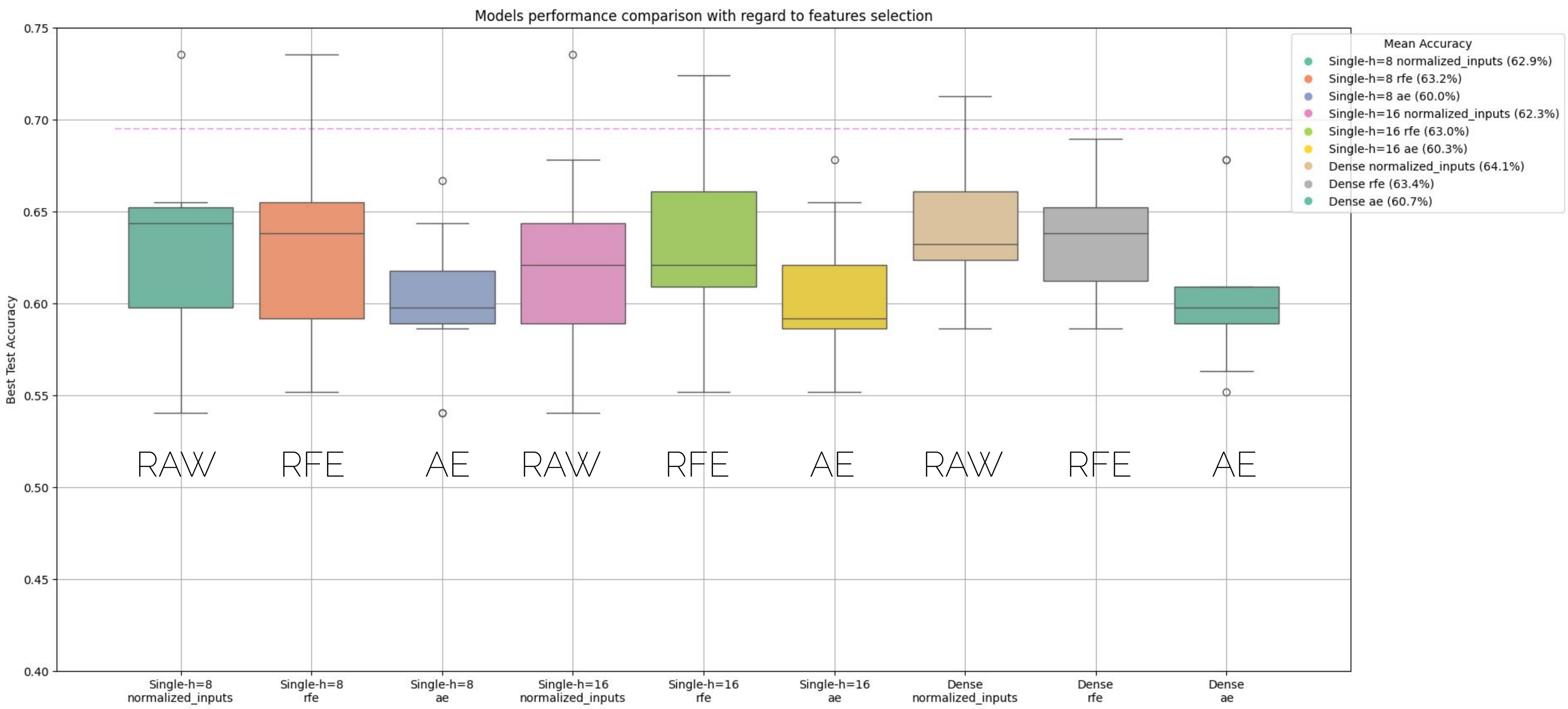
⟶ BILATERAL FILTER / NON LOCAL MEANS *



[*]Antoni Buades, Bartomeu Coll, and Jean-Michel Morel, *Non-Local Means Denoising*, Image Processing On Line, 1 (2011), pp. 208–212.

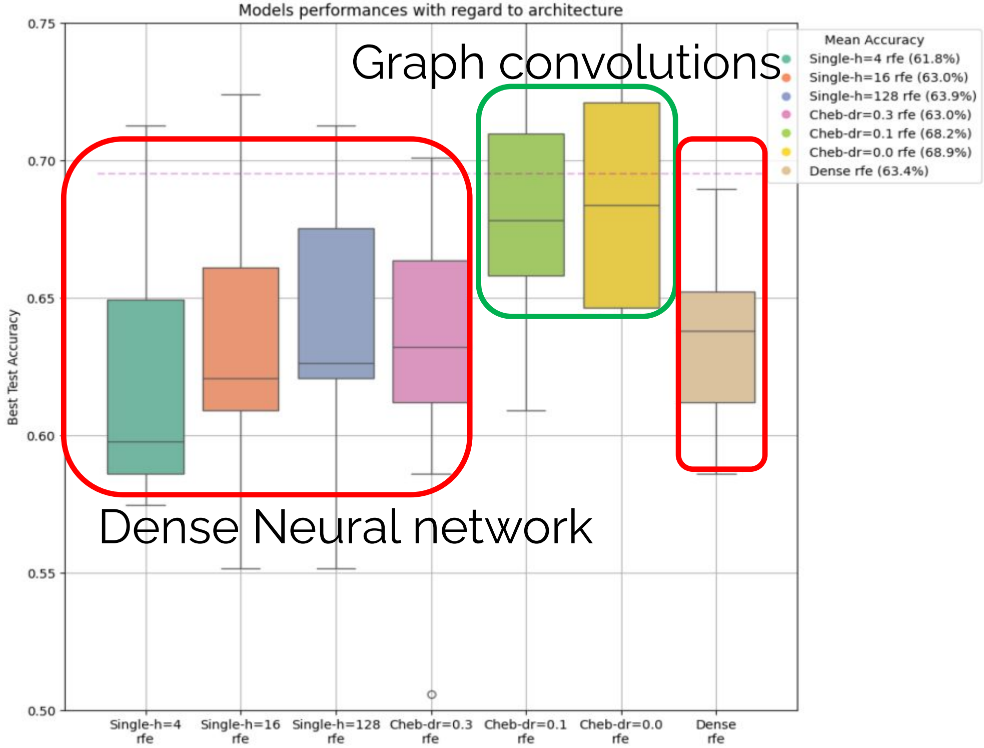
Results

- Relevance of reduce feature dimension?
  - **NOT REALLY**



Results

- Relevance of using graph?
  - **YES** ~ +5% accuracy



Extra questions

A the beginning of the project, we asked ourselves a few questions on the paper:

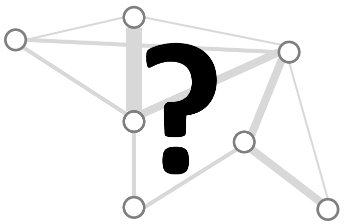
- Does the graph structure bring anything?
  -> yes
- What does feature reduction do?
  -> nothing significant
- Would adding metadata to the input feature help a dense NN do better?
- How does the research world process MRI?
- Can we think of a toy example?
  -> not trivial, focus on ABIDE dataset

Medical domain
Expectations V.S. Reality

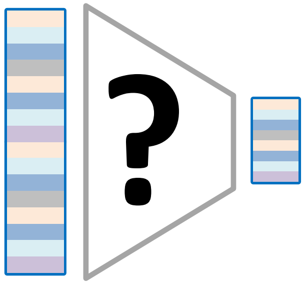
Expectation	Reality
f-MRI 4D volumes	Time series
Graph, big memory requirement	GPU laptop Nvidia T500 4Gb
« Spectral graph convolutions »	Message passing

Investigation

- Is the use of graphs truly relevant?

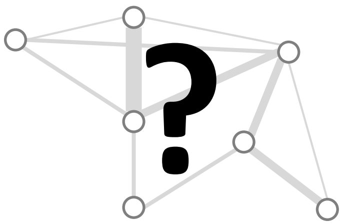


- Is input dimension reduction relevant?



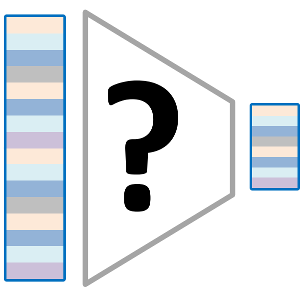
Investigation

- Is the use of graphs truly relevant?
    - Dense Neural Networks
    - Graph Convolutional Networks **+5%**



- Is input dimension reduction relevant?

No significant differences between
- Raw input features
- Recursive Feature Elimination
- Auto Encoder's Latent vector



Overview

- Paper summary - Overview
  - Node
  - Edges
  - GCN
- Our contributions
  - Graph relevance
  - Feature dimension reduction
- Discussion

Medical domain
Expectations V.S. Reality

Expectation	Reality
MRI 3D volumes ?	Time series
Huge medical dataset	861 patients – curse of dimensionality
Graph, big memory requirement	Nvidia T500 4Gb ~ enough
« Spectral graph convolutions »	Message passing, analogy NL-means