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

MATHÉMATIQUES
VISION
APPRENTISSAGE

Paper overview

Disease prediction
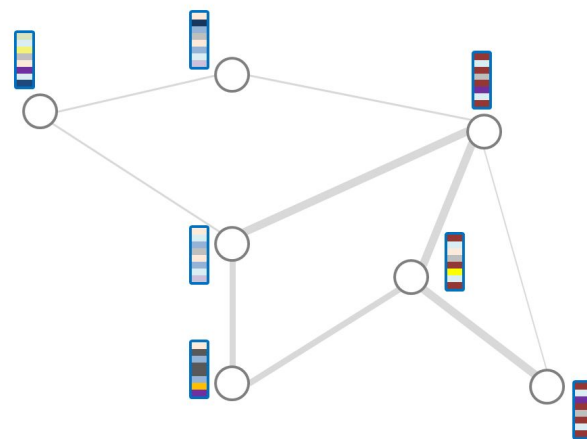
Contribution: Graph Convolution Networks [*]

Datasets : ABIDE → Autism | ADNI → Alzheimer

Population graph

○ Nodes = Patients

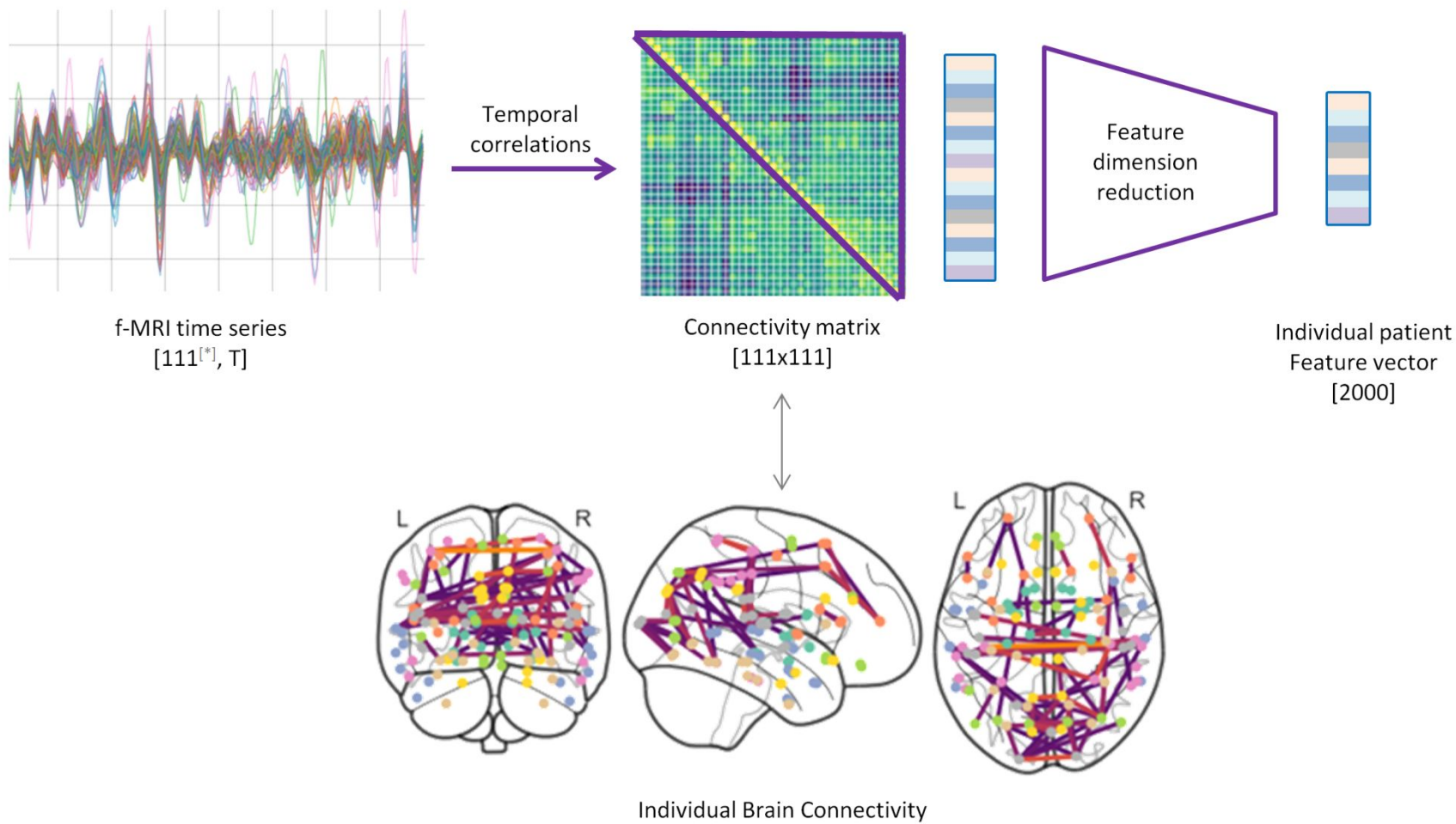╱ Edges = Similarity between patients

▯ Feature vector

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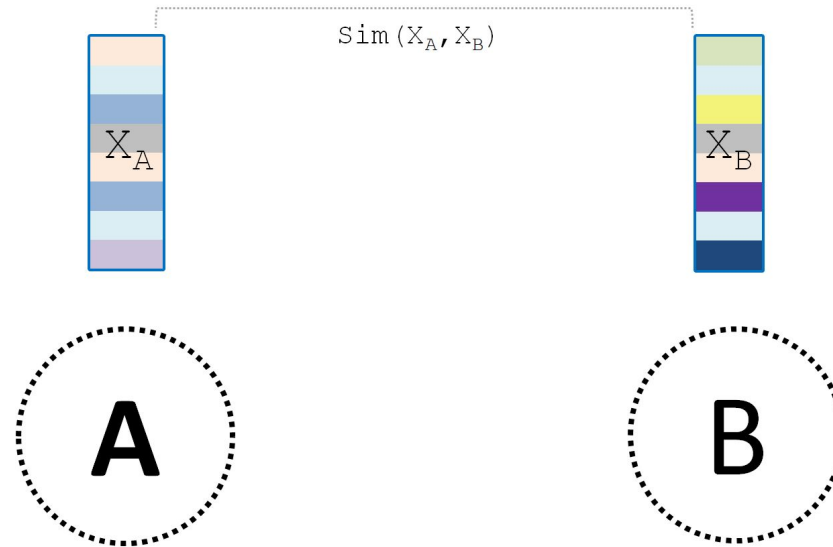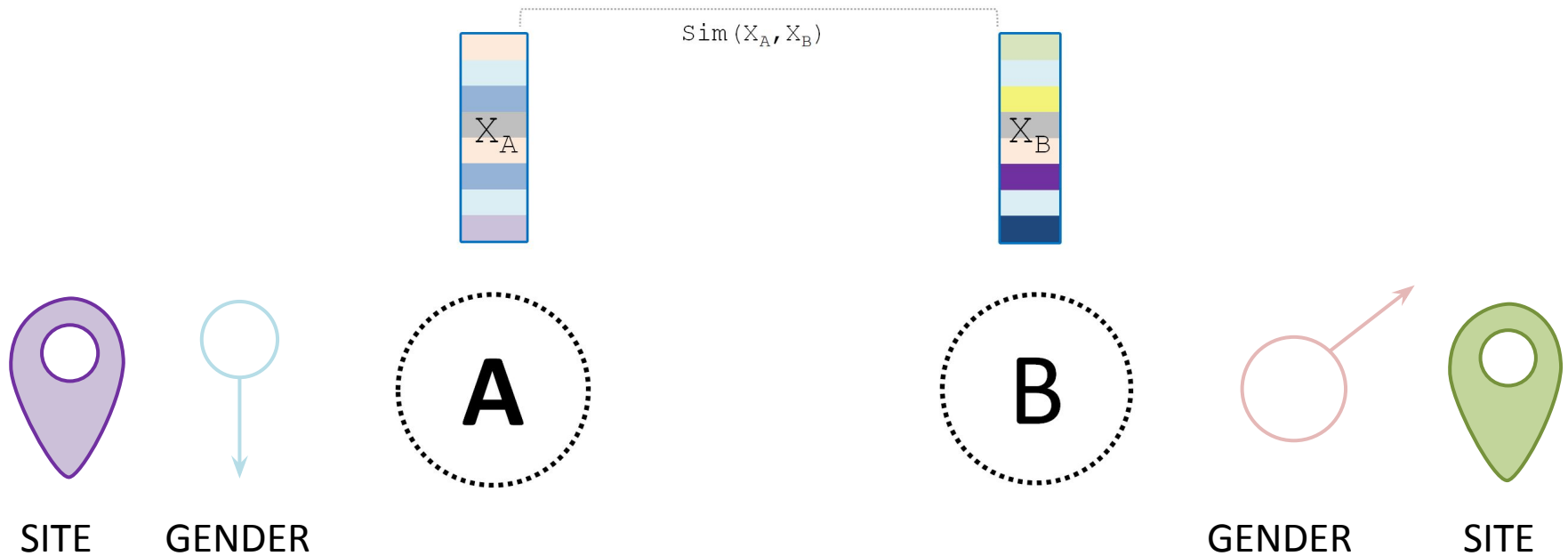
Defining graph nodes

1 patient ↔ 1 node

f-MRI time series
[111[*], T]

Temporal correlations

Connectivity matrix
[111x111]

Feature dimension reduction

Individual patient
Feature vector
[2000]

Individual Brain Connectivity

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

MATHÉMATIQUES VISION APPRENTISSAGE

Defining graph edges

$$\texttt{Sim}(X_A, X_B)$$

X_A

X_B

A

B

Defining graph edges

$$\texttt{Sim}(\texttt{X}_A, \texttt{X}_B)$$

SITE GENDER A B GENDER SITE

Defining graph edges

$$\mathrm{Sim}(X_A, X_B)$$

X_A X_B

A ⎯⎯⎯ B

SITE   GENDER   GENDER   SITE

x1

MATHÉMATIQUES
VISION
APPRENTISSAGE

Defining graph edges

$\text{Sim}(X_A, X_B)$

X_A

X_B

SITE   GENDER   A ——— B   GENDER   SITE

x2
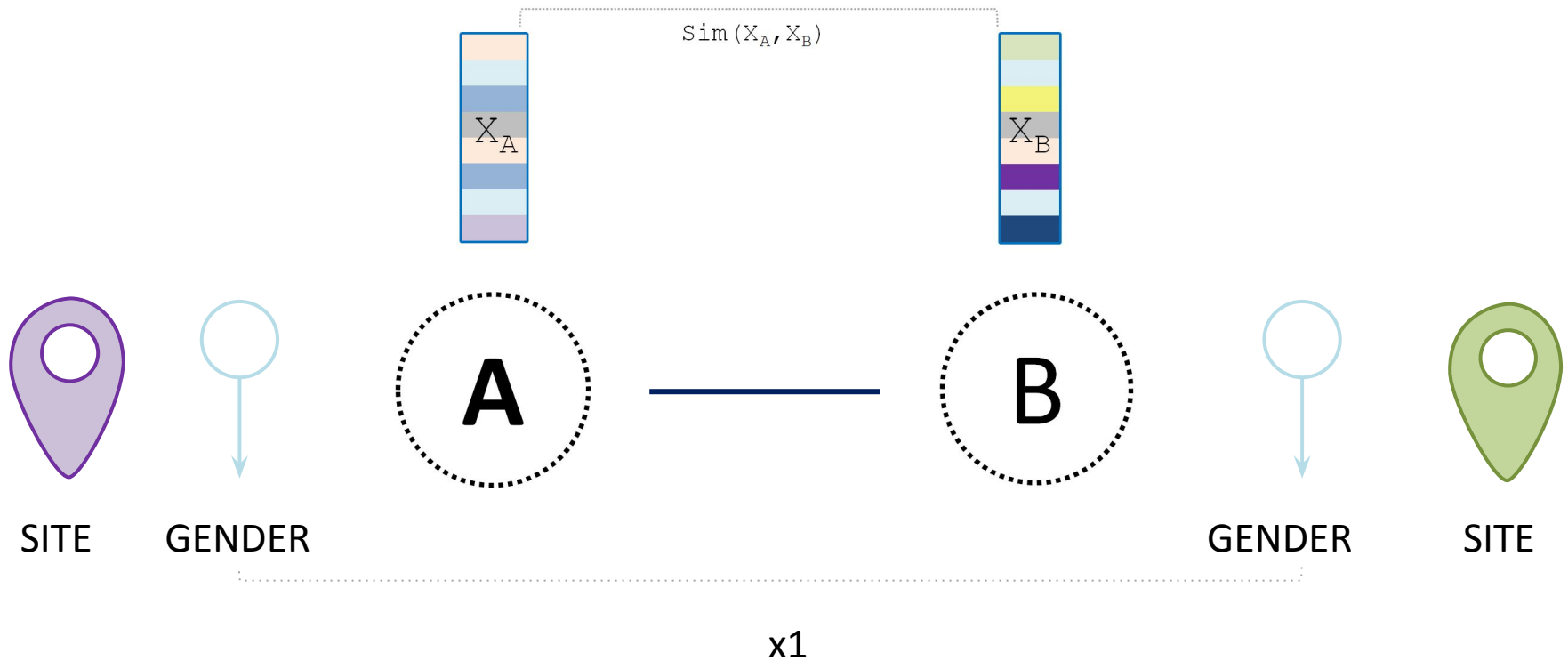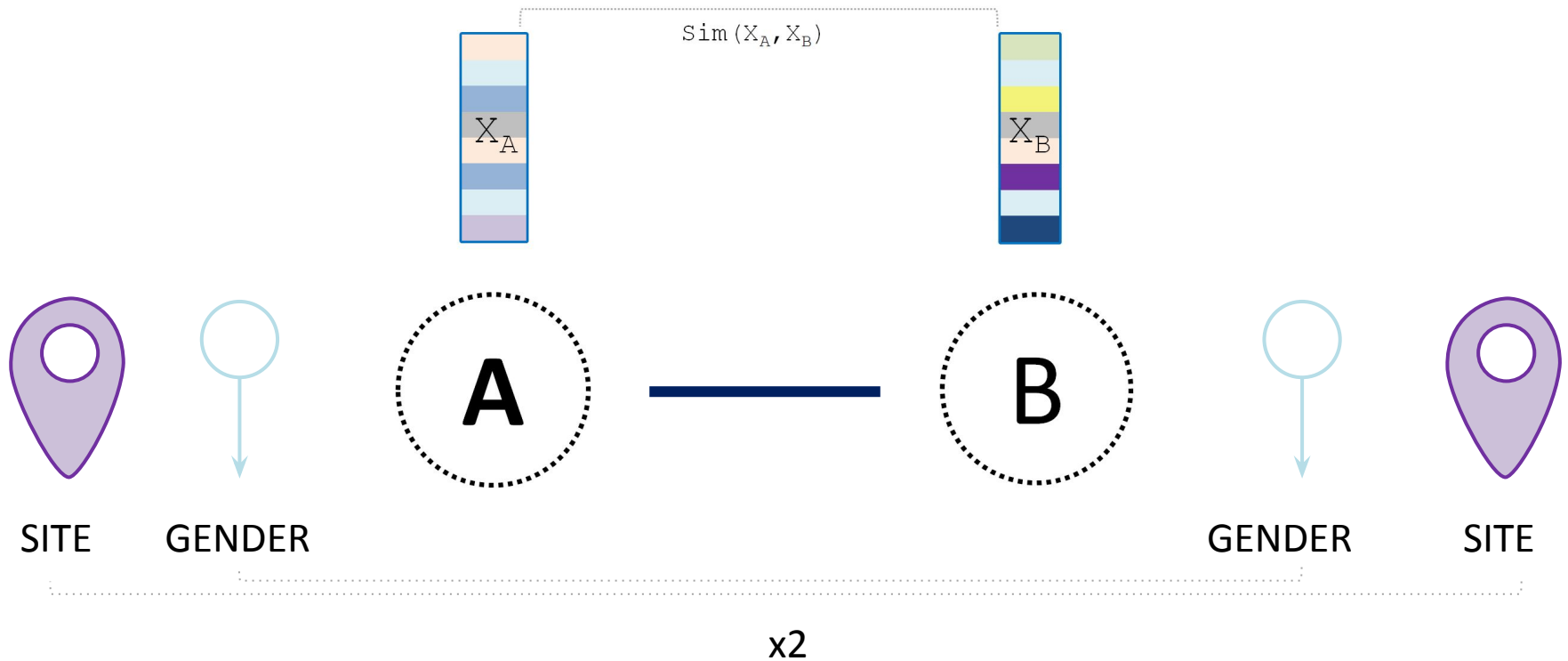
Defining graph edges

Graph Convolution Network
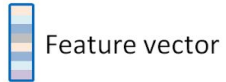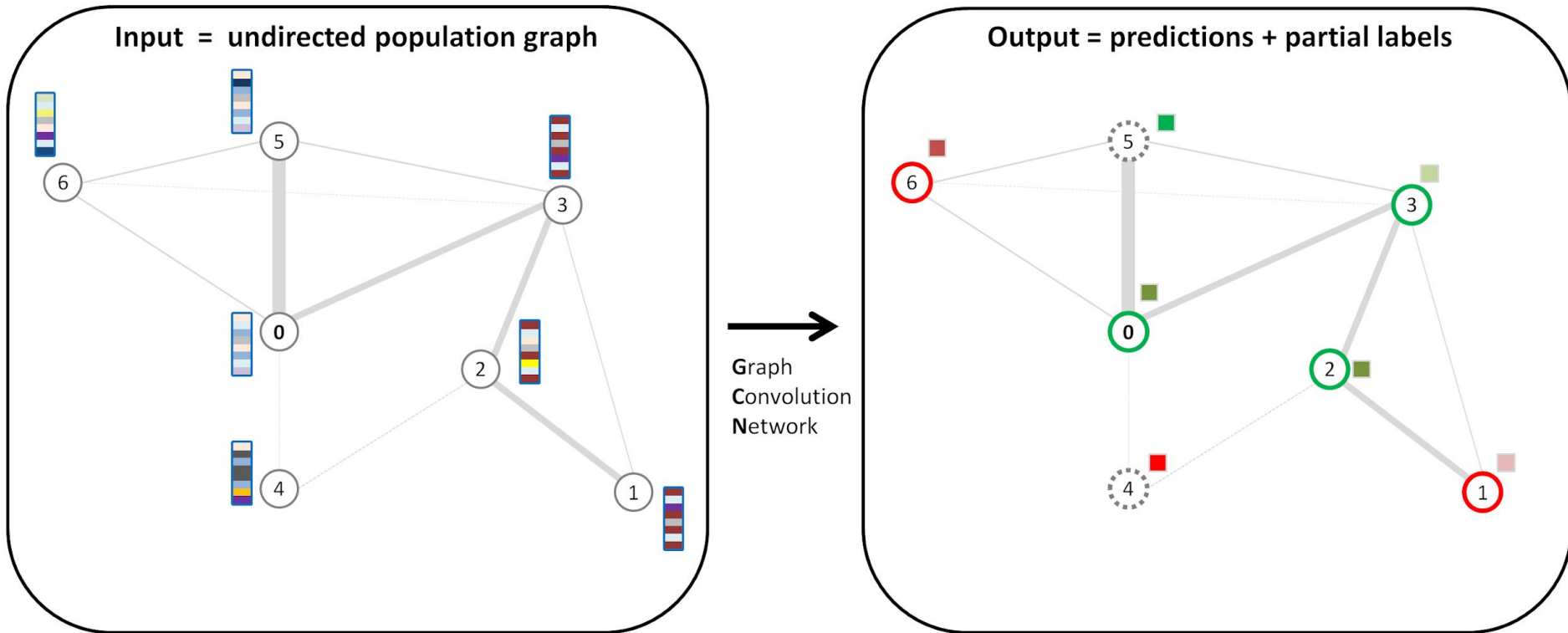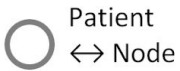
Graph Convolution **Network**

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

Θ

N
patients

=861

d h

=6000

X^T
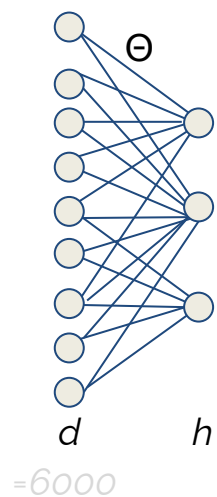
d *=6000*
features
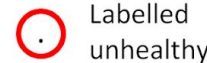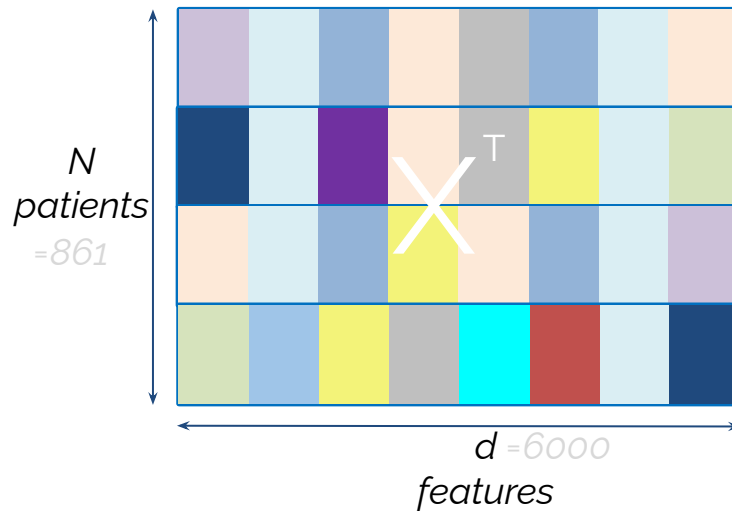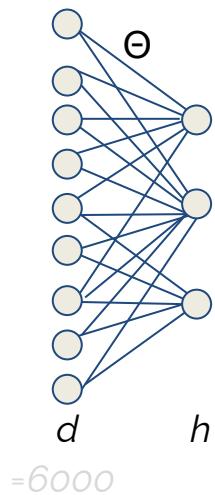
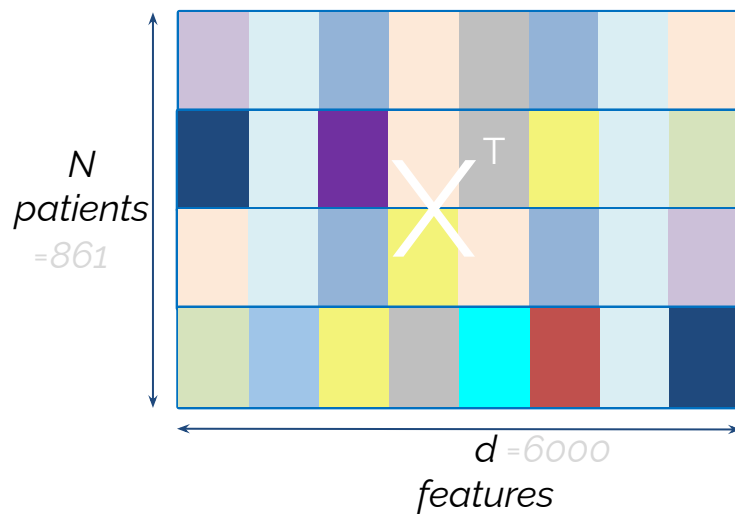Fully connected *[h,d]*   Input features *[N, d]*

Graph Convolution Network

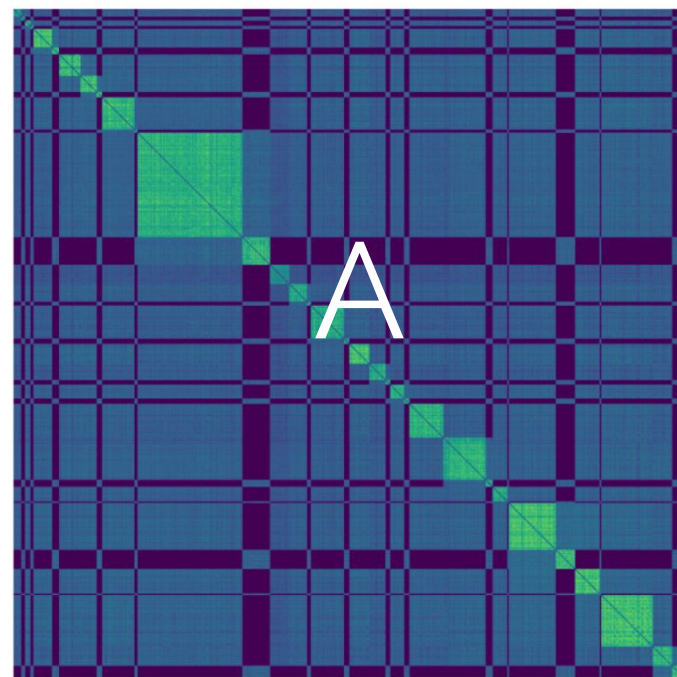$$\underbrace{Y}_{h,N} = \underbrace{\Theta}_{h,d} \; \underbrace{X}_{N,d}^{T} \; \underbrace{A}_{N,N}$$

Basic message passing

A

N patients
=861

d =6000
features

d =6000

h

Θ

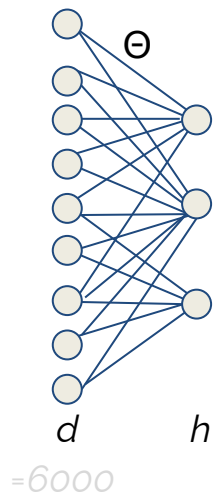Fully connected *[h,d]*

Input features *[N, d]*
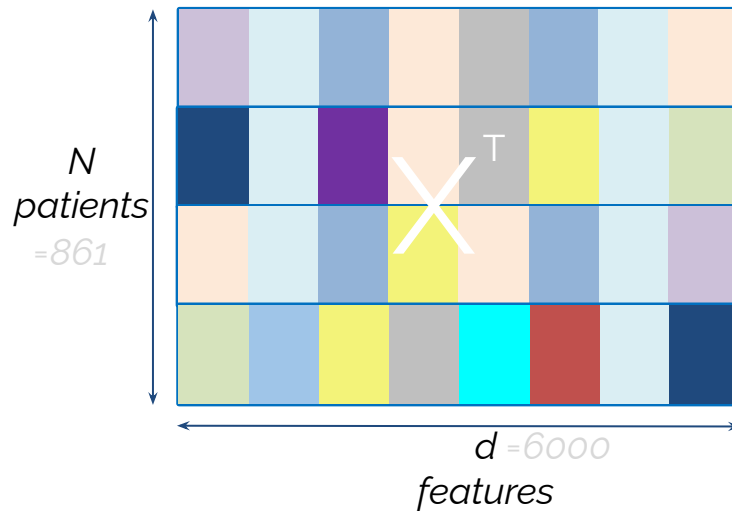
Adjacency matrix *[N, N]*

Analogy: Non Local means - denoise similar colors on images

Graph Convolution Network

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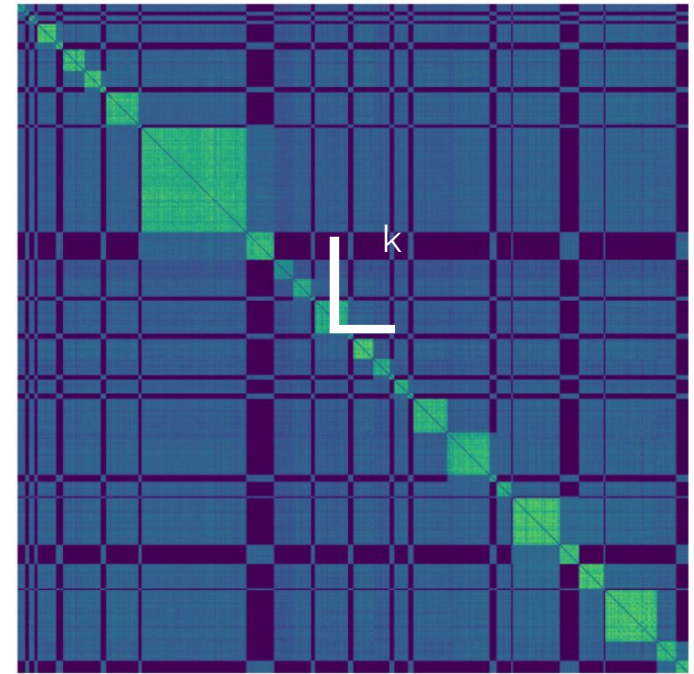
Basic message passing

Fully connected *[h,d]*

Input features *[N, d]*
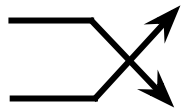
Power of Laplacian matrix *[N, N]*

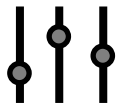*Variant in the paper : ChebConv - polynomial of the Laplacian L=I-A ➡ gives access to a larger neighborhood*

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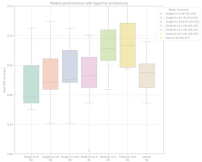
Accuracy 68.9% +/- 4% on ABIDE (Autism Disorder Spectrum)
Ability to reproduce results (~69.5%)
→ Compare various models

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
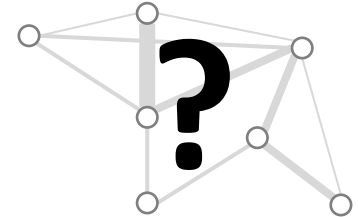
Reproducible results

Code from scratch with documentation.

MATHÉMATIQUES
VISION
APPRENTISSAGE

Results

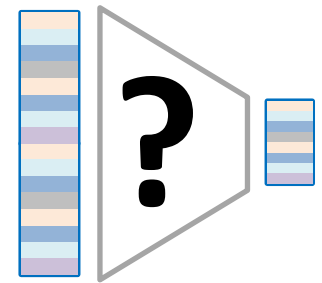- Is the use of graphs truly relevant?
  Compare
  - Dense Neural Networks
  - Graph Convolutional Networks **+5%**

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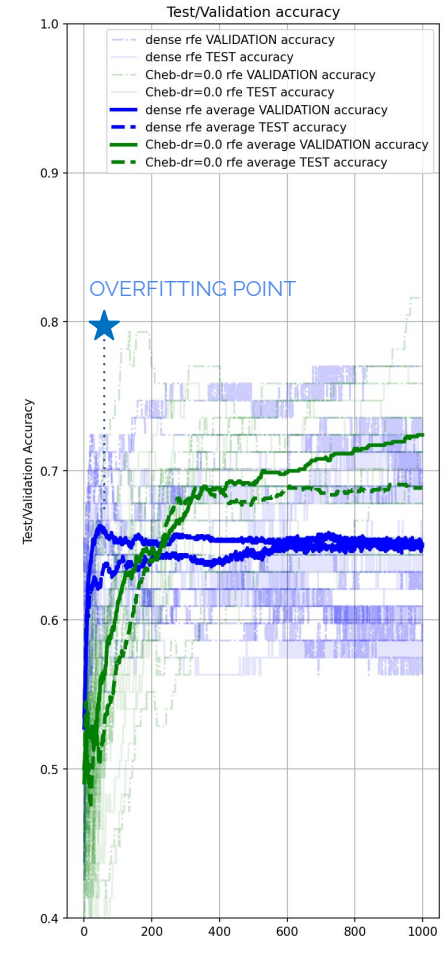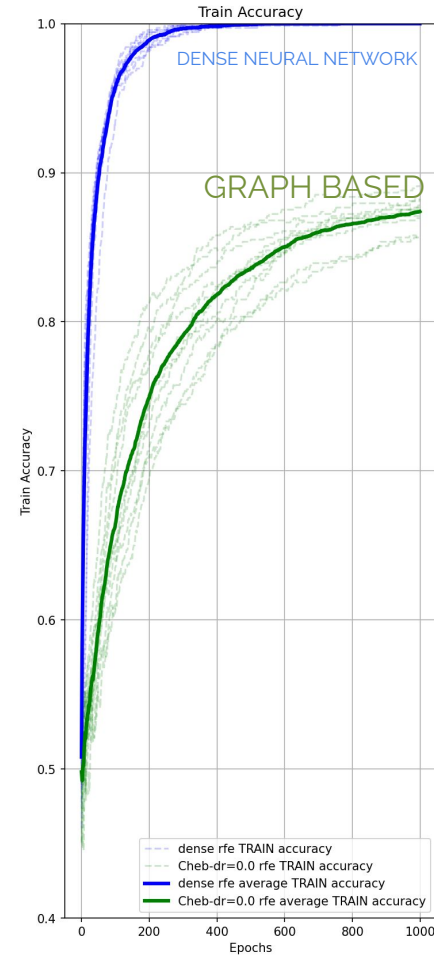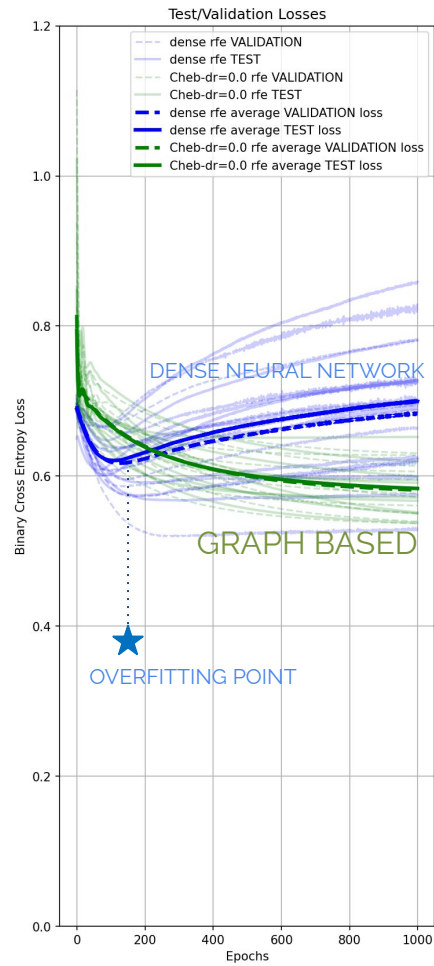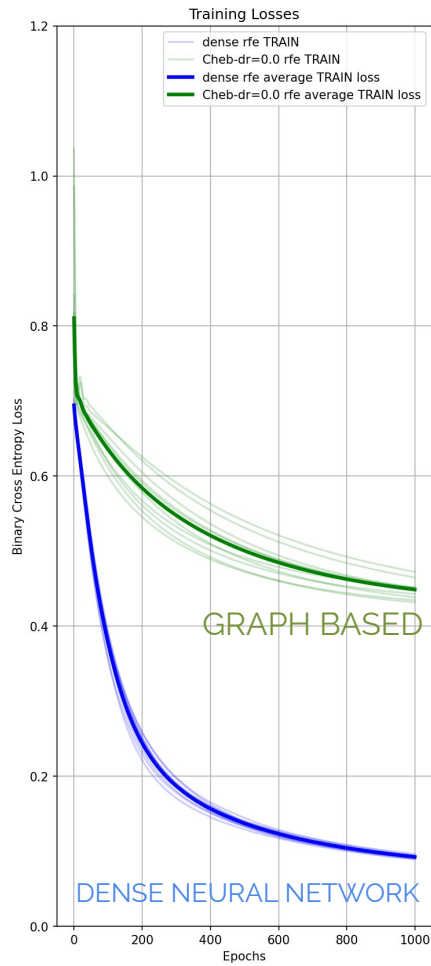 - Auto Encoders' Latent vector **-3%**
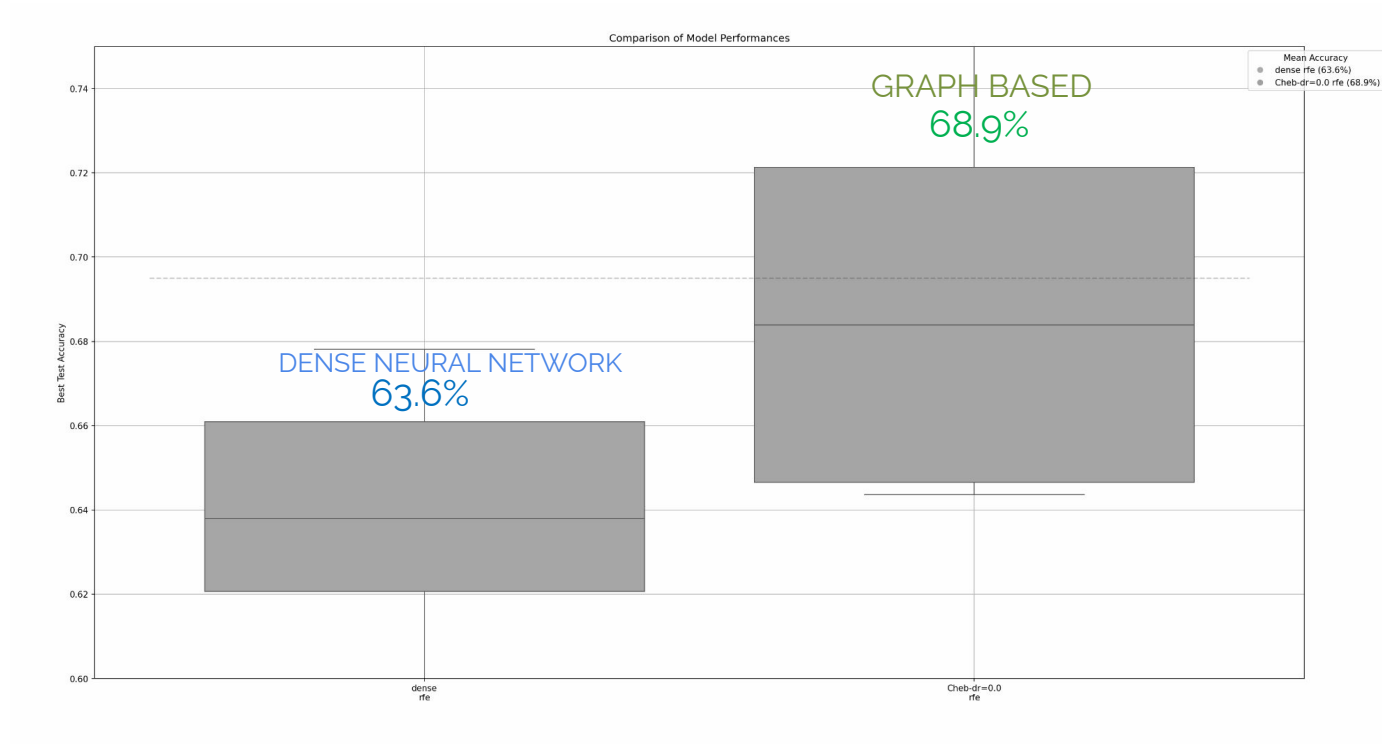
Conclusion

- 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 »

MATHÉMATIQUES
VISION
APPRENTISSAGE
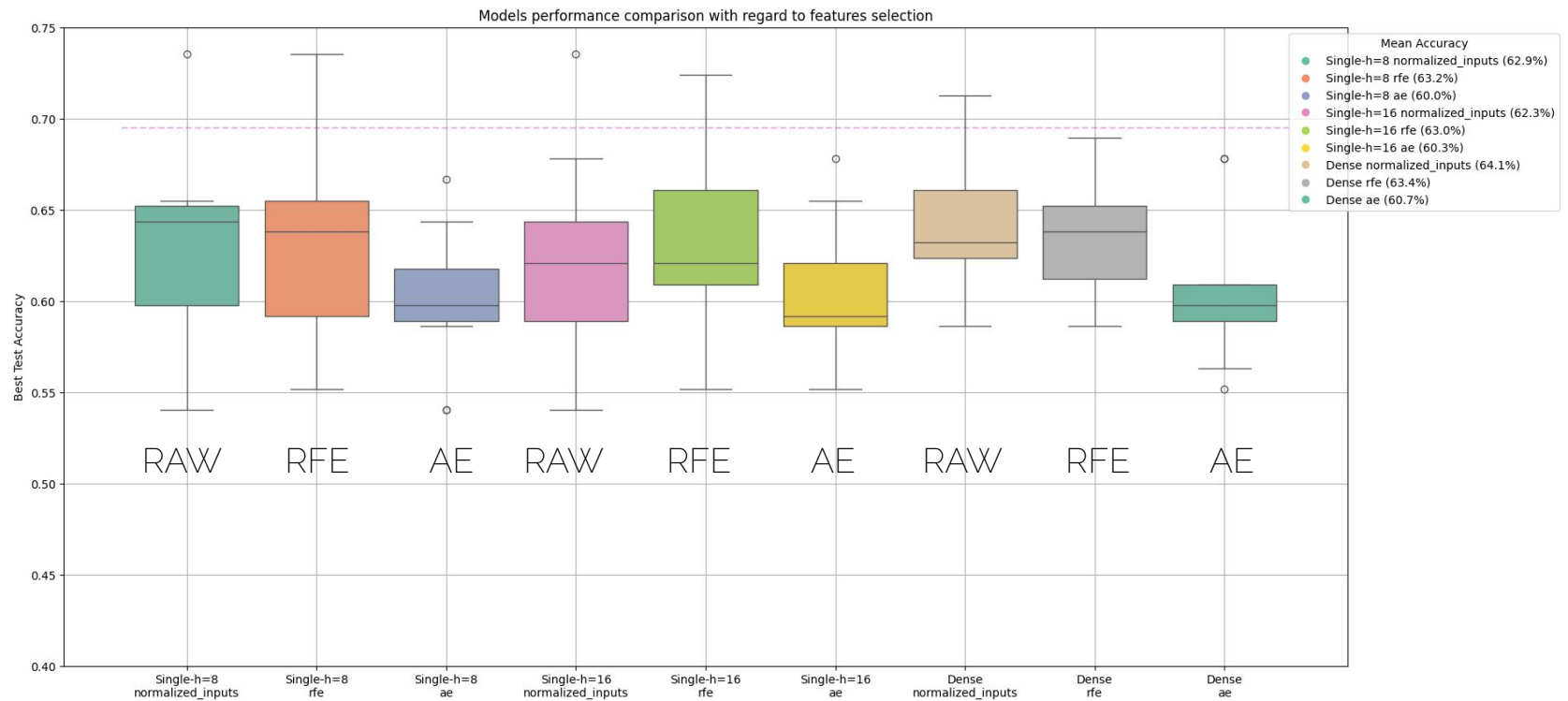
Thank you

Training curves

Dense Neural Network VS Graph Convolution
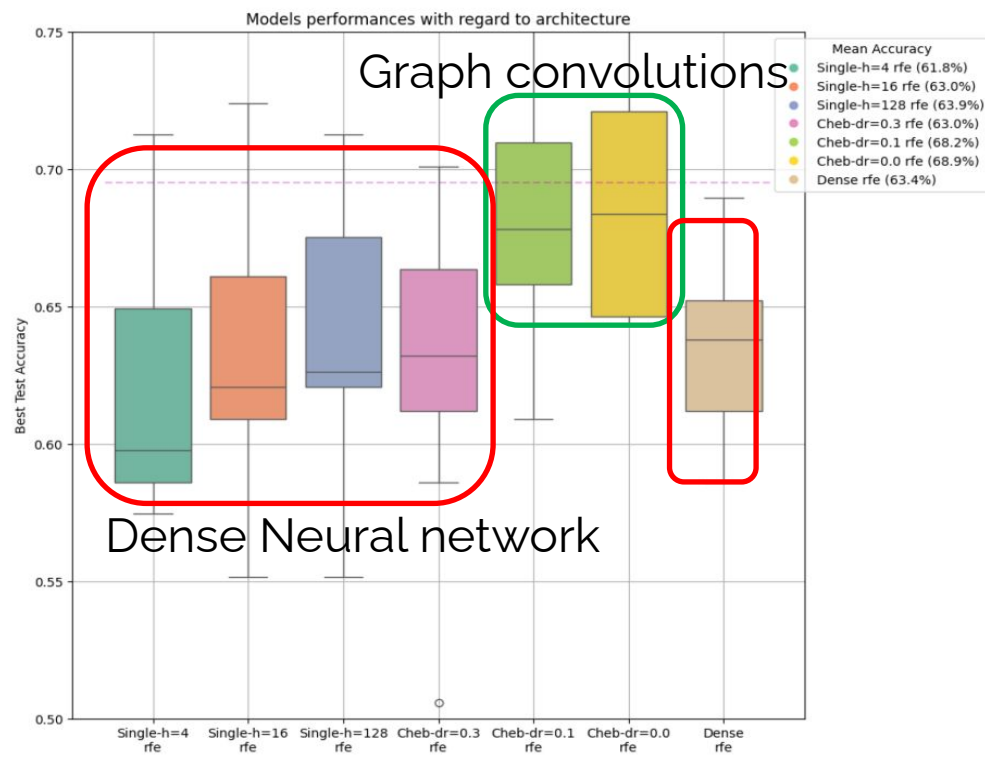
Results

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

Results

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

Medical domain
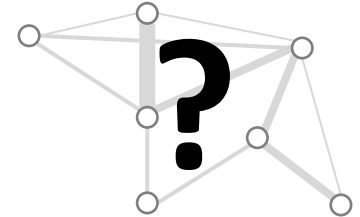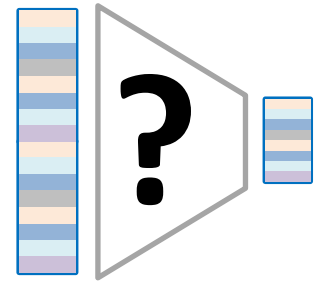# Expectations V.S. Reality

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

MATHÉMATIQUES
VISION
APPRENTISSAGE

Investigation

- Is the use of graphs truly relevant?

- Is input dimension reduction relevant?

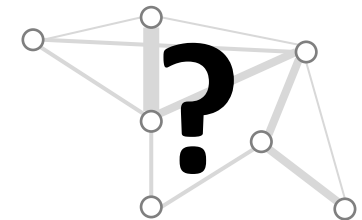MATHÉMATIQUES VISION APPRENTISSAGE
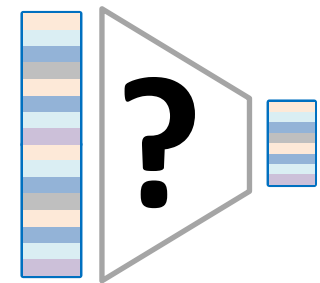
Investigation

- 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

MATHÉMATIQUES
VISION
APPRENTISSAGE

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