Empirical survey on graph classification methods for Brain Networks

Faculty of Information Engineering, Computer Science and Statistics

Corso di Laurea in Data Science

Candidate

Ilaria Taddei
ID number 1853733

Thesis Advisor	Co-Advisor
Prof. advisor	Dr. co-advisor

Academic Year 2020/2021

Thesis defended on 22 September 2015
in front of a Board of Examiners composed by:

Prof. ... (chairman)

Prof. ...

Prof. ...

Prof. ...

Prof. ...

Prof. ...

Prof. ...

––

Empirical survey on graph classification methods for Brain Networks
Bachelor's thesis. Sapienza – University of Rome

This thesis has been typeset by LaTeX and the Sapthesis class.

Author's email: ilariataddei.95@gmail.com

Abstract

Our brain, as our entire body, is a perfect as complicated machine. For this reason is important to analyse it in an effective and efficient way. An important role in this purpose is given to machine learning and similar algorithmic methods. In fact, many papers have been written on different methods to study the brain as graph or as matrix. Most of all, in those papers they want to recognize if a person is affected by a mental decease, such as, for example, Autism or Schizophrenia.

In this work, the aim is to study and compare some of those methods. First of all is important to have in mind which are the basic elements with which the papers work, such as graphs, classification and brain connectomes. Then, the algorithms taken in consideration can be classified according to the technique used, like Neural Networks or features embeddings. For each method many experiments are made, to see how they work, in which case are more useful, and how good are the results. The experiments are made with different datasets, to see how the methods adapt and how general they are.

Future works can be implemented, studying more methods and maybe finding a method that could be easily used by neuroscientists.

Contents

Chapter 1

Introduction

I n the frame of astronomical spectroscopy...

In Chapter 2 we briefly present...

In Chapter 3 we summarize...

1.1   Motivations

1.2   Background

To understand better all the methods, we should do some refresh of what we will encounter during the lecture. The main arguments we should consider are **Graphs**, **Classification** and **Connectomes**.

Graphs

There are many subjects in which graphs are used, and consequently many definitions. In mathematics, a graph is a structured set of objects that, in pairs, could have a relation between themselves. Each object is called a *vertex*, and each relation between a pair of vertices is called an *edge*. So, a graph $G(V, E)$ is a pair where V are the vertices and E are the edges. Graphs are represented in diagram form, where vertices are in circle form and edges are lines that join pairs of vertices with a relation.
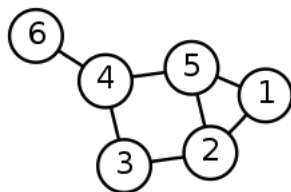

Figure 1.1

Graph could be *directed* or *undirected*. In a direct graph, edges have an orientation, the links between vertices can be represented by arrows going from one vertex to the other. In an edge (x, y) directed from x to y, the vertices are called respectively *tail* and *head* of the edge.
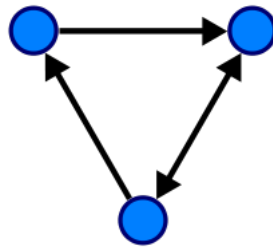
Figure 1.2

1.3 Related works

1.4 Outline

Chapter 2

Survey

2.1 General

2.2 All methods

We can divide the methods in the several techniques that the papers use to build the graph classification: Neural Networks, Embedding, Deep matrix factorization, Graph Fourier transform, Graph kernels, and Network classification.

2.2.1 Neural Networks

GroupINN: Grouping-based Interpretable Neural Network for Classification of Limited, Noisy Brain Data

This paper proposes a grouping-based interpretable neural network model, GroupINN, that classifies cognitive performance with 85% fewer parameters than baseline deep models, while also identifying the most predictive brain subnetworks within several task-specific contexts. In the design of the neural network is included the idea of node grouping. In this way the model learns the node grouping and extracts the graph features jointly.

The problem statement of this paper is: given a set of subjects, each with corresponding fMRI data and a label associated with a certain phenotype, we seek to devise an efficient, interpretable, and parsimonious neural network model that can predict each phenotype with high accuracy.

To reduce the number of parameters used in the model, they adopted the idea of multi-graph clustering (where the goal is to find a common clustering across multiple graphs) to summarize the original graph into a supergraph with each cluster as a supernode.

The neural network is formed by three different types of layers: node grouping layer, graph convolutional layer and fully connected layer. The node grouping layer is designed to "hide" the non-indicative ('noisy') edges by grouping them into a cluster, thus highlighting the indicative edges: two nodes are assigned to different groups if their connection is identified as important. Graph convolutional layers are used to capture the structure of the supergraph.

The neural network is also divided in two branches, one processes the positive graphs and one the negative ones. All in all, the architecture consists of three kinds of layers and two branches. The input graph is the correlation matrix W. The first layer is a dimensionality reduction layer and the output is a matrix W^s representing the

supergraph. Following the dimensionality reduction layer, three graph convolutional layers are used. At last, the positive and negative outputs of the previous layer are concatenated, flattened and sent to the fully connected layer (with softmax activation).

Functional Brain Network Classification for Alzheimer's Disease Detection with Deep Features and Extreme Learning Machine

2.2.2 Embedding

Explainable Classification of Brain Networks via Contrast Subgraphs

In this paper they introduce an approach for classifying brain networks based on extracting contrast subgraphs, i.e., a set of vertices whose induced subgraphs are dense in one class of graphs and sparse in the other. The model is extremely simple, with just one parameter, excellent interpretability and good classification accuracy. They exploited contrast subgraph TD-ASD (the subgraph that maximizes the difference between the number of edges in the class TD, typically developed brains, with respect to the same measure for class ASD, people with autism spectrum deseas). Vertex size represents the importance of the vertex in discriminating the two classes. In fact they ended up with some important rules: 1. If an individual exhibits more than 62 edges among the 15 vertices of the contrast subgraph ASD-TD, then there are high chances that the individual is affected by ASD; 2. If the number of edges induced by the contrast subgraph ASD-TD is smaller than half of the number of edges induced by the contrast subgraph TD-ASD, then there are high chances that the individual is not affected by ASD. 3. If the number of edges induced by the contrast subgraph ASD-TD is smaller than the number of edges induced by the contrast subgraph TD-ASD, then there are high chances that the individual is affected by ASD.
Each individual is represented by an undirected unweighted graph with $|V| = 116$ vertices.

Stable Biomarker Identification For Predicting Schizophrenia in the Human Connectome

Unsupervised Network Embedding for Graph Visualization, Clustering and Classification

Supervised classification of structural brain networks reveals gender differences

2.2.3 Deep matrix factorization

Multi-modality disease modeling via collective deep matrix factorization

2.2.4 Graph Fourier transform

Integration Of Network Topological Features And Graph Fourier Transform For Fmri Data Analysis

2.2.5   Graph kernels

Sub-network Kernels for Measuring Similarity of Brain Connectivity Networks in Disease Diagnosis

2.2.6   Network classification

Network Classification With Applications To Brain Connectomics

Chapter 3

Experiments

3.1 Experimental setup

With datasets explanation

3.2 Results

(Data Science)

3.3 Discussion

Chapter 4

Conclusions

4.1 Future works

Bibliography

[1] Blanche P.A., Gailly P., et al., "*Volume phase holographic gratings: large size and high diffraction efficiency*", Optical Engineering, Vol. 43, No.11, November 2004

[2] Cirasuolo M., et al., *"MOONS Science Report"*, MOONS Document Number: VLT-TRE-MON-14620-0001, Issue: 1.0, 31st January 2013

[3] European Southern Observatory, http://www.eso.org