

# Learning from Graphs

A workshop proposal for The Web Conference (WWW'21)

Michaël Defferrard

2020-12-04

## Organizer

I am a Machine Learning researcher, currently pursuing a PhD at the École Polytechnique Fédérale de Lausanne (EPFL) with Prof. Pierre Vandergheynst. My main research interest is the modeling, analysis, and understanding of structured data. Data structured by networks, such as brain activity supported by neural connections, or manifolds, such as the temperature and wind fields on the Earth. To this end, I am developing Deep Learning to exploit that structure, and thrive to solve real problems. Besides research, I am teaching Data Science and Machine Learning (with graphs).

More details online: [website](#), [CV](#), [twitter](#), [google scholar](#), [github](#).

## Abstract

A graph encodes relations between objects, such as distances between points or hyperlinks between websites. You will learn how to extract information about that relational structure. This information is crucial to characterize an object through its local connectivity or an entire graph through its global connectivity. On top of that structure, a network may possess data about the objects or the relations, such as a point's color or an hyperlink's click-through rate. You will learn how to leverage a graph to analyze this data. Leveraging the structure that underlies data is an important concept, from physical symmetries dictating conservation laws to the efficiency of convolutional neural networks. The tutorial will walk you from the basics to some of the latest developments, with an emphasis on intuitions and working knowledge.

## Topic and relevance

Graphs are ubiquitous among topics covered by The Web Conference: they organize the web both physically and logically, they model social interactions, they power recommender systems. The theory and tools I will develop in the workshop are however agnostic to the particular application. They can be used whenever data is structured by some sort of relation, with applications ranging from the design of proteins to the inference of cosmological parameters from sky observations.

A rich and mature theory with tools has been developed over decades in the fields of Graph Theory and Network Science. The versatility of graphs to represent relational data and exploit their symmetries has recently caught the attention of the Signal Processing and Machine Learning communities. This tutorial will walk participants from established knowledge to some of the latest developments, following a logical path that will touch adjacent topics, such as calculus and group theory, to build useful connections.

## Duration

A full-day (6-7 hours) is required to cover the breadth of the topic and to allow enough time for the participants to experiment.

## Interaction style

Eight periods of 45 minutes made of 15 minutes of theory with slides followed by 30 minutes of practice with Jupyter notebooks. Q&A and discussions of theoretical or practical aspects anytime.

## Intended audience and level

**Intended audience:** people with theoretical or practical interests about graphs and data on graphs. Adapted to both researchers and practitioners.

**Prerequisite knowledge:** scientific python programming, basic linear algebra, no prior knowledge about networks is necessary.

**Learning objectives:** Participants will learn how to identify network data, how to deal with it, and what can be learned from it. They will know the basics of information processing over networks, and how to devise a machine learning system based on network data. Finally, the hands-on experience will give them the confidence to apply those tools in practice.

## Previous editions

I taught tutorials on similar topics at the Applied Machine Learning Days ([slides](#), [github](#), [conference](#)) and at the GraphSIP summer school ([github](#), [summer school](#)). I managed and co-taught a course at EPFL for four years ([2019](#), [2018](#), [2017](#), [2016](#)). I also maintain the [PyGSP](#), a python package for graph signal processing, which contains tutorials and examples.

While topics are similar, the material matures over time and is adapted to the interests of the venue's audience.

## Tutorial materials

PDF slides and jupyter notebooks on github (to be run either locally if the computer was prepared with a given [conda](#) environment or in the cloud with [binder](#)).

## Online format

- Theory with slides presented live or pre-recorded.
- Continuous instant messaging (I guess there will be a platform for the whole conference).
- One-on-one Q&A and discussions over “break-out rooms”.
- Code can be run in the cloud to avoid installation issues.

## Additional info for hands-on tutorials

- Tutorial duration: 6 hours.
- OS and tools: any OS with [conda](#) or a web browser.
- Software licenses: none.
- Setup instructions: `git clone xxx` (or zip download) then `conda create -f environment.yml`. Can work in the cloud with a browser if that fails. Detailed instructions will be available on the github repository (like [here](#)).

## Video snippets

Not of teaching, but some research talks are on [youtube](#).