From algebraic topology to medicine: topological data analysis and physiological signals

Supervisors

- Dr. Brian Tervil: Centre Borelli (ENS Paris Saclay/Université Paris Saclay, CNRS, Université de Paris, SSA, INSERM). brian.tervil@ens-paris-saclay.fr
- Prof. Laurent Oudre: Centre Borelli (ENS Paris Saclay/Université Paris Saclay, CNRS, Université de Paris, SSA, INSERM). laurent.oudre@ens-paris-saclay.fr

Keywords

Topological data analysis, signal processing, machine learning, behavioral neuroscience, medicine.

Our laboratory

This internship will take place at the Centre Borelli (ENS Paris Saclay/Université Paris Saclay, CNRS, Université de Paris, SSA, INSERM). The Centre Borelli is a multidisciplinary research unit focusing on all applications of mathematics, neuroscience and biomedical research. It brings together multidisciplinary teams of mathematicians and experts in medicine, physics, mechanics, biology, and engineering, to conduct research actions motivated by real data and use cases from science and industry. It offers courses of excellence in applied mathematics, neurosciences, ergonomy and bioengineering on promising themes with high technological and societal stakes.

The societal stakes are considerable and offer many opportunities to the student who will choose this project in the field of the interface of engineering sciences, computer science, new technologies and medicine... In case of success, the internship may be the subject not only of scientific publications but also of a valorization project.

Environment

The trainee will have the opportunity to be fully integrated into an interdisciplinary team of mathematicians, clinicians, neuroscientists, statistical modelers and computer scientists. The internship will take place at ENS Paris Saclay (Gif sur Yvette) and Université de Paris (Paris). The trainee will attend weekly seminars of the MLMDA team. The duration of the internship is between 4 and 6 months and the trainee will be paid according to the standard internship wage (around 600 euros per month).

Motivation

Mathematical context

Topological data analysis (TDA) [1] has been, for the past twenty years, a great source of novel perspectives in data science, and has proven its potential in various areas such as GPS trajectory reconstruction [2], object recognition [3], and neuroscience [4]. Finding its roots in the theory of algebraic topology, developed at the end of the nineteenth century by Henri Poincaré, TDA offers a description of properties that are inherent to the "shape" of a given dataset. These properties are often represented in the form of so-called *persistence diagrams*, which provide *a summary of the evolution of the arrangement of data at several scales or levels*. Thanks to their underlying theoretical guarantees, especially their robustness to noise, TDA's topological invariants often constitute a much suitable information in many situations. While TDA has been applied for several years to some subdisciplines of neuroscience, especially the study of the anatomic organization of the central nervous system, its potential in behavioral neuroscience, neurophysiology, or in medicine, has only been recently exploited [5, 6].

Clinical context

Multiple sclerosis (MS) is a demyelinating disease of the central nervous system with varying clinical presentations and progression. Gait, and more generally lower limb impairment, are considered the most important sources of disability among MS patients. While semi-quantitative indicators exist and provide clinical scales for the evaluation of the disease evolution, the latter are not adapted to inter-individual variability, nor permit a fine movement quantification. For a few years, light, affordable wearable inertial measurement units (IMU) have been providing to hospital practitioners and researchers the possibility to refine these scales, and study at a so far unreachable level of detail the intrinsic characteristics of motor degeneracy in this disease [7, 8, 9].

From mathematics to neurology

The signals coming from IMU describe, within the case of motor function, the dynamics of the human body, in the form of time series. The tools from TDA provide a suitable framework for studying the periodicity, or in other words the regularity of this dynamics [10]. The main idea of the proposed project arises from the above argument. More precisely, the internship will consist of analyzing, through their underlying topological structure, the regularity of gait signals among healthy and MS subjects. The relevance of the topological information for longitudinal follow up of gait quality among MS patients will be studied, under the hypothesis that gait degeneracy can be detected via variations of its regularity. The correlation between topological variations in signals and variations in pre-existing indicators will then be investigated [8, 9].

This study is part of a broader research program on the contribution of concepts from geometry and topology to behavioral neuroscience and medicine, especially for longitudinal follow up and multimodal sensorimotor analysis, and more generally for the quantification of normal and pathological human behavior.

Objectives and internship progress

The goal of this internship is to become familiar with certain concepts from TDA, and apply them to the study of physiological signals. More precisely, it will consist of evaluating the relevance of the topological information for the longitudinal follow up of gait in MS patients. First, the student will familiarize with:

1. The concepts and tools from TDA, as well as the GUDHI Python library [11].
2. A database consisting of IMU signals from gait trials of healthy and MS subjects, acquired by the Centre Borelli through the collaboration between mathematicians, engineers and neurologists.
3. A series of papers forming the main publications of the Centre Borelli in the topic, especially [7, 8, 9].

Second, the student will have to implement and test functions for:

4. Evaluating one or several constructions of persistence diagrams from time series.
5. Evaluating distances between persistence diagrams corresponding to different gait trials.
6. Comparing the aforementioned distances with pre-existing indicators of MS clinical evolution.

A report about the discussed concepts and obtained results will be submitted at the end of the internship in the form of a thesis, and will be the subject of a defense.

Preliminary planning

Months 1 and 2

- Bibliography and familiarization with the theoretical concepts.
- Familiarization with the database.

- Familiarization with the GUDHI library and the functions relevant to the study.

Month 3

- Evaluation of several construction techniques of persistence diagrams from time series and choice of a method suitable to the study.
- Evaluation of distances between persistence diagrams for different gait trials.

Month 4

- Comparison of the obtained trajectories of persistence diagrams with the ones from pre-existing indicators.
- Thesis writing.

Required skills

- First year master's degree level in mathematics.
- Familiarity with TDA concepts (MVA course recommended).
- Basic knowledge in machine learning.
- Knowledge of Python and its main libraries.

References

[1] Chazal, Frédéric & Michel, Bertrand. (2017). An introduction to Topological Data Analysis: fundamental and practical aspects for data scientists.

[2] Dey, Tamal & Wang, Jiayuan & Wang, Yusu. (2018). Graph Reconstruction by Discrete Morse Theory. 10.4230/LIPIcs.SoCG.2018.31.

[3] Li, Chunyuan & Ovsjanikov, Maks & Chazal, Frédéric. (2014). Persistence-Based Structural Recognition. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 10.1109/CVPR.2014.257.

[4] Lida, Kanari & Dlotko, Pawel & Scolamiero, Martina & Levi, Ran & Shillcock, Julian & Hess, Kathryn & Markram, Henry. (2016). Quantifying topological invariants of neuronal morphologies.

[5] Dindin, Meryll & Umeda, Yuhei & Chazal, Frédéric. (2019). Topological Data Analysis for Arrhythmia Detection through Modular Neural Networks.

[6] Rossi-deVries, Jasmine & Pedoia, Valentina & Samaan, Michael & Ferguson, Adam & Souza, Richard & Majumdar, Sharmila. (2018). Using Multidimensional Topological Data Analysis to Identify Traits of Hip OA. Journal of Magnetic Resonance Imaging. 48. 10.1002/jmri.26029.

[7] Oudre, Laurent & Barrois-Müller, Rémi & Moreau, Thomas & Truong, Charles & Vienne-Jumeau, Aliénor & Ricard, Damien & Vayatis, Nicolas & Vidal, Pierre-Paul. (2018). Template-Based Step Detection with Inertial Measurement Units. Sensors. 18. 4033. 10.3390/s18114033.

[8] Vienne-Jumeau A, Oudre L, Moreau A, Quijoux F, Vidal PP, Ricard D. (2019). Comparing Gait Trials with Greedy Template Matching. Sensors (Basel) .19(14):3089. Published 2019 Jul 12. doi:10.3390/s19143089

[9] Vienne-Jumeau, Aliénor & Oudre, Laurent & Moreau, Albane & Quijoux, Flavien & Edmond, Sébastien & Dandrieux, Mélanie & Legendre, Eva & Vidal, Pierre & Ricard, Damien. (2020). Personalized Template-Based Step Detection From Inertial Measurement Units Signals in Multiple Sclerosis. Frontiers in Neurology. 11. 10.3389/fneur.2020.00261.

[10] Ravishanker, Nalini & Chen, Renjie. (2019). Topological Data Analysis (TDA) for Time Series.

[11] Maria, Clément & Boissonnat, Jean-Daniel & Glisse, Marc & Yvinec, Mariette. (2014). The Gudhi Library: Simplicial Complexes and Persistent Homology. 10.1007/978-3-662-44199-2_28.