

Groupe de Travail : Topological Data Analysis (TDA)

Rubén Martos & Vincent Runge

1 Sujet

Le but du groupe de travail est de donner une introduction au sujet de l'*Analyse Topologique des Données*. Nous expliquerons les fondements mathématiques de la TDA (tels que l'homologie persistante) ainsi que son intégration dans l'apprentissage automatique.

En quelques mots, la TDA vise à utiliser les invariants de la topologie algébrique pour classer les données et pour les comparer. Ainsi, la TDA part du principe que les *données ont une forme*. Les techniques de la TDA sont orientées à découvrir la géométrie sous-jacente. Parmi les nombreuses applications concrètes, nous pouvons citer la reconnaissance de formes et l'analyse d'images qui ont permis de classer des bactéries et des neurones ou d'identifier certaines formes de cancer.

2 Infos pratiques

- **Lieu des exposés** : salle des séminaires du 4ème étage de l'IBGBI (LaMME)
- **Horaire** : 14h – 16h (sauf contre-indication)

3 Programme

1. Introduction : 15 mai 2025

- Motivation et contexte de la TDA dans l'apprentissage automatique.
- Invariants topologiques : caractéristique d'Euler et nombres de Betti.
- Homologie simpliciale.

2. Inférence topologique : 21 mai 2025

- t-épaissements et le théorème du nerf : le complexe de Čech.

- Homologie persistante.
 - Décomposition des modules de persistance : code-barres et le diagramme de persistance.
 - Le théorème de stabilité.
3. **Représentations de la persistance pour l'apprentissage automatique : 28 mai 2025**
- Courbes de Betti.
 - Paysages de persistance.
 - Images de persistance.
 - Métriques de persistance et techniques-noyaux.
 - La courbe d'Euler.
4. **Aspects statistiques de l'homologie persistante : 11 juin 2025**
- Régions de confiance d'un diagramme de persistance.
 - Analyse statistique d'une population de diagrammes de persistance.
 - Intégration de la TDA dans l'apprentissage automatique.
 - Applications représentatives.
5. **Clôture : 18 juin 2025**
- Défis computationnels, questions ouvertes et perspectives de recherche.
 - Séminaire de recherche: “*Contributions à l'inférence de l'homologie persistante et applications*” par Hugo Henneuse (Université Paris Saclay - LMO et INRIA Saclay - Datashape)

4 Bibliographie

1. Bobrowski Omer and Skraba Primoz. A universal null-distribution for topological data analysis. *Scientific Reports*, 13, 12274 (2023).
2. Chazal, Frédéric and Michel Bertrand. An introduction to topological data analysis: fundamental and practical aspects for data scientists. *Frontiers in artificial intelligence*, 4 (2021).
3. Edelsbrunner Herbert and Harer John L. Computational topology: An introduction. *American Mathematical Society* (2010).
4. Hacquard Olympio and Lebovici Vadim. Euler characteristic tools for topological data analysis. *Journal of Machine Learning Research*, 25:1-39, (2024).
5. Wasserman Larry. Topological data analysis. *Annual Review of Statistics and its Application*, 5:501-532 (2017).