TOPOLOGICAL DATA ANALYSIS
COURSE SYLLABUS

INSTRUCTOR: BRIAN HEPLER

Course Overview:

This course introduces the mathematical foundations and computational methods of Topological Data Analysis (TDA), with emphasis on persistent homology and its applications. Students will learn to extract robust topological features from complex datasets, implement TDA algorithms, and apply these techniques to real-world problems in data science, machine learning, and emerging fields like quantum computing. The course balances rigorous mathematical theory with hands-on computational practice using modern Python libraries.

Lecture Topics

(1) Introduction to Topological Data Analysis

(2) Simplicial Complexes and Constructions

(3) Simplicial Homology

(4) Persistent Homology

(5) Algorithms and Computational Aspects

(6) Stability and Statistical Methods

(7) TDA in Machine Learning

(8) Advanced Topics and Applications

Prerequisites:

- **Point-Set Topology:** Open/closed sets, continuity, compactness, connectedness, metric spaces, homeomorphisms
- **Linear Algebra:** Vector spaces, matrices, eigenvalues, linear transformations
- **Multivariable Calculus:** Partial derivatives, gradients, optimization basics
- **Programming:** Basic Python proficiencies
- **No prior knowledge of algebraic topology required** (we will build this from scratch)

Course Materials

Primary References.

- Carlsson, G. "Topology and Data." Bulletin of the American Mathematical Society 46.2 (2009): 255-308
- Edelsbrunner, H. & Harer, J. "Computational Topology: An Introduction"
- Ghrist, R. "Elementary Applied Topology"
- Hatcher, A. "Algebraic Topology" (selected chapters)

Software Tools.

- **GUDHI** (Geometry Understanding in Higher Dimensions)
- **Ripser** (Efficient Vietoris-Rips computation)
- **Scikit-TDA** (Machine learning integration)
- **Giotto-TDA** (Topological machine learning)

LECTURE SCHEDULE

Lecture 1: Introduction to Topological Data Analysis.

- Motivation: Why topology for data analysis?
- Challenges in high-dimensional data analysis
- Overview of the TDA pipeline
- From point clouds to topological spaces
- Introduction to simplicial complexes
- Case studies: sensor networks, biological data

Interactive Component:

- Visualizing point clouds and their topological features
- First look at persistence using simple examples

Lecture 2: Simplicial Complexes and Constructions.

- Simplicial complexes: formal definitions and examples
- Čech, Vietoris-Rips, and Alpha complexes
- Witness complexes for large datasets
- The Nerve Theorem and its applications
- Computational trade-offs between different constructions

Interactive Component:

- Building and visualizing different complex types
- Comparing computational complexity and accuracy

Lecture 3: Simplicial Homology.

- Chains, cycles, and boundaries
- Boundary operators and chain complexes
- Homology groups and Betti numbers
- Geometric interpretation: components, holes, voids
- Computing homology through linear algebra

Interactive Component:

- Step-by-step homology computation examples
- Implementing basic homology calculations

Lecture 4: Persistent Homology.

- Motivation: why persistence matters
- Filtrations and persistence modules
- Birth and death of topological features
- Persistence diagrams and barcodes
- The algebraic structure of persistence

Interactive Component:

- Computing persistence for real datasets
- Visualizing and interpreting barcodes

Lecture 5: Algorithms and Computational Aspects.

- Matrix reduction algorithms
- Standard vs. dual algorithms
- Computational complexity and optimizations
- Software implementations comparison
- Handling large-scale datasets

Interactive Component:

- Benchmarking different algorithms
- Performance optimization techniques

Lecture 6: Stability and Statistical Methods.

- Stability theorem for persistence diagrams
- Bottleneck and Wasserstein distances
- Statistical inference on persistence diagrams
- Persistence landscapes and images
- Hypothesis testing and confidence intervals

Interactive Component:

- Analyzing noisy data and stability
- Statistical comparisons of persistence diagrams

Lecture 7: TDA in Machine Learning.

- TDA features for machine learning pipelines
- Persistence kernels and vectorization methods
- Topological layers in neural networks
- Mapper algorithm and data visualization
- Case studies in classification and clustering

Interactive Component:

- End-to-end ML pipeline with TDA features
- Implementing topological layers in neural networks

Lecture 8: Advanced Topics and Applications.

- Multiparameter persistence
- TDA in quantum computing and quantum data
- Persistent cohomology and circular coordinates
- Current research frontiers
- Open problems and future directions

Final Project Overview:

- Students select a dataset and perform complete TDA analysis
- Options include: genomics, neuroscience, materials science, quantum systems

Learning Outcomes

By the end of this course, students will be able to:

- Construct and analyze simplicial complexes from data
- Compute and interpret persistent homology
- Apply TDA methods to real-world datasets
- Integrate topological features into machine learning pipelines
- Understand the theoretical foundations and stability guarantees of TDA
- Implement efficient algorithms for topological computations
- Critically evaluate when and how to apply TDA methods