Syllabus and other details for Course 59300-TDA: Topological Data Analysis

Course Details :
Course 59300-TDA

Instructor: Tamal K. DEY

Time: TTh 1:30–2:45 pm

https://www.cs.purdue.edu/homes/tamaldey/course/CTDA/CTDA.html

Theme. Topological Data Analysis has played a synergistic role in bringing together research work from computational geometry, algebraic topology, data analysis, and many other related data-centered application areas. In recent years, the field has undergone particular growth in the area of data science. The application of topological techniques to traditional data analysis, which before has mostly developed on a statistical setting, has opened up new opportunities. This course is intended to cover this aspect of topological data analysis along with the developments of generic techniques for various topology-centered problems. *Objectives* of this course are:

1. Be familiar with basics in topology that are useful for computing with data
2. Master a subset of algorithms for computing: Betti number, topological persistence, homology cycles, Reeb graphs, discrete Morse structures, multiparameter persistence
3. Be familiar with how to design algorithms for problems inTopological Data Analysis has played a synergistic role in bringing together research work from computational geometry, algebraic topology, data analysis, and many other related data-centered application areas. In recent years, the field has undergone particular growth in the area of data science. The application of topological techniques to traditional data analysis, which before has mostly developed on a statistical setting, has opened up new opportunities. This course is intended to cover this aspect of topological data analysis along with the developments of generic techniques for various topology-centered problems. applications dealing with data
4. Be familiar with how to research the background of a topic in topological data analysis, machine learning

Textbooks, materials.

- Computational Topology and Data Analysis, T. K. Dey and Y. Wang, Cambridge U. Press.
- Computational Topology, Herbert Edelsbrunner and John L. Harer, AMS.
- Class materials and notes posted on the web-page for the course: Taken mostly from the book: Computational Topology and Data Analysis, T. K. Dey and Y. Wang, Cambridge U. Press.

Grading.

- Assignment or term paper: 30%
- Midterm: 30%
- Final: 40%

Topics.

1. Basics of Topology:
    - Topological spaces, metric space topology [Chapters from book 1 and 2] [Notes]
    - Maps: homeomorphisms, homotopy equivalence, isotopy [Notes]
    - Manifolds [Notes]
2. Complexes on data, Algorithms:
    - Simplicial complexes [Munkres book][Notes]
    - Chech complexes, Vietoris-Rips complexes [Notes]
    - Witness complexes [deSilva-Carlsson04 paper][Notes]
    - Graph induced complexes [DeyFanWang13 paper][Notes]
3. Homology:
    - Chains, boundaries, homology groups, betti numbers [Notes, Munkres book]
    - Induced maps among homology groups [Notes, Munkres book]
    - Relative homology groups [Notes, Munkres book]
    - Local homology groups [Notes, Munkres book]
    - Cohomology groups [Notes, Edelesbrunner and Munkres book]
4. Topological persistence, Algorithms:
    - Filtrations, Persistent homology [Notes, C-VII Edelsbrunner-Harer book]
    - Persistence algorithm [Notes, C-VII Edelsbrunner-Harer book]
    - Persistence diagram [Notes, C-VIII Edelsbrunner-Harer H book]
    - Variations on persistence algorithm [Notes, CarlssonSilvaMorozov09 paper on zigzag persistence, DeyFanWang13SM paper on cohomology persistence]
5. General Persistence (Zigzag), Algorithms:
    - Persistence modules from simplicial maps [Notes]
    - Zigzag persistence [Notes]
    - Level Set persistence [Notes]
    - Extended persistence [Notes]
    - Nerves, Mapper and Multiscale Mapper
    - Nerves [Notes]

- Mapper [Notes]
- Multiscale Mapper [Notes]

Other possible topics:

6. Topology inference from data:
   - Computing homology from data [Notes, ChazalOudot08 paper on homology inference, CCGGO09 paper on interleaving of persistence modules]
   - Sparsification to handle big data [Presentation slides], [Sheehy12 paper on sparsified Rips complex, DeyFanWang13 paper on subsampling]
7. Computing optimized homology cycles:
   - Computing shortest basis cycles on surfaces [Notes, EricksonWhittlesey05 paper on greedy basis construction]
   - Computing shortest basis cycles from data points [Notes, DeySunWang09 paper on shortest basis from point data]
   - Localizing a cycle class [Notes, DeyHiraniKrishnamoorthy10 paper on LP algorithm for shortest homologous cycle]
8. Persistence on graphs and Reeb graphs:
   - Reeb graphs, Algorithms [Notes]
   - Interleaving distance [Notes]
   - Comparing graphs with persistence summary [Notes]
9. Multiparameter persistence:
   - Multiparameter persistence [Notes]
   - Computing indecomposables [Notes]
   - Computing distances among persistence modules [Notes]

Instructor: **Tamal K. Dey**.
Classes: **TTH 1:30–2:45 pm**
Office hours: **TTH 3:00-4:00 pm. or by appointment**
Grading Policy: **Midterm Exam:** 30%**, Term paper or Assignment:** 30%**, Final** 40%
Course web-page

`https://www.cs.purdue.edu/homes/tamaldey/course/CTDA/CTDA.html`