CS 6170: Computational Topology, Spring 2019
Project 3 (Bonus Project)
Topological Data Analysis for Data Scientists

Dr. Bei Wang

School of Computing
Scientific Computing and Imaging Institute (SCI)
University of Utah
www.sci.utah.edu/~beiwang
beiwang@sci.utah.edu

April 4, 2019

Project 3 (Bonus Project) Overview

Playing with Kepler Mapper: Part I

- Posting date: 4/4/2019. Due date: 4/18/2019.
- Set up: Install Kepler Mapper
  `https://kepler-mapper.scikit-tda.org/`
- Part 1: Play with build-in examples under *examples* folder (5 points).
  - Modify *horse* example so that the overlapping parameter is changed to 50%. (1 points)
  - Modify *horse* example by replacing the mapper clustering method with K-means clustering from scikit-lean, with parameter $n\_clusters$ set to 2 or 3. (2 points)
  - Modify *makecircles* example by replacing the lens to a different projection direction; and explain what this modification mean to the data (2 points)

Playing with Kepler Mapper: Part 2

- Part 2: Using Kepler Mapper with a high-dimensional example (15 points)
  - Prepare Boston Housing data for Kepler Mapper, using at least 5 dimensions (5 point)
    - `http://math.furman.edu/~dcs/courses/math47/R/library/mlbench/html/BostonHousing.html`
    - You can download the dataset in other places (as long as you specify where your data is coming from and include the processed data in the submission)
  - Use "tax" as a filter function, using 40% overlap, show the mapper results. What kind of insights can you get from the exploratory process? You should specify what other parameters you choose for the mapper algorithm. (5 points)
  - Use both "tax" and "crim" as filter functions (2D mapper, see *examples/breast-cancer* for an example of how to create 2D mapper, show the mapper results. Again, specify other parameters for the mapper algorithm. What kind of insights can you get from the exploratory process?(5 points)

Submission

- Processed data files (csv), Screen shots (PDF) and source code (python and html files) should be included as part of the submission.
- Project report (PDF) and source code should be submitted in a single ZIP file.