Using topological autoencoders for global and local topology

Filip Cornell

KTH Royal Institute of Technology

Introduction

Using the Mapper method to visualize high-dimensional data **should capture the topological properties of the data**. Today, there are few or no guarantees that the filtering functions used for the Mapper does this.

We propose:

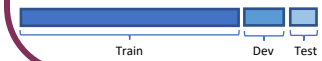
- Using topological autoencoders (TAEs)[1] to capture local and global topologies, constructing *more representative Mapper graphs*.
- *TAEs* capture the shape of the data through a *Persistent Homology-regularized loss.*
- Comparing with other filtering functions on high-dimensional manifolds to validate mapper graphs.

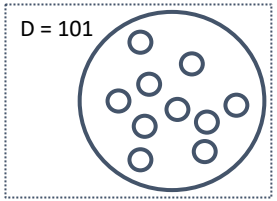
Data

Our data consists of 11 100-dimensional hyperspheres originally constructed by Moor et al.[1]

- 10,000 datapoints.
- Points sampled uniformly at random on surfaces of different spheres.

Splits





Method

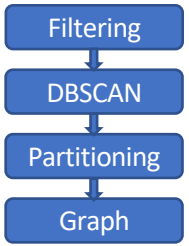
Empirical investigation
Compare 6 different functions for Mapper over different settings.

Grid search over:

- Overlap fraction 0.025 – 0.4 (step of 0.025).
- Number of intervals: 5– 45 (step of 5).

Metric
Average amount of spheres in each node in the Mapper graph.

Mapper pipeline

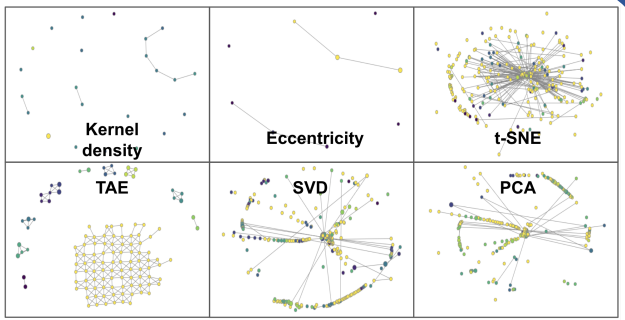


Results

- TAE reaches **optimal separation of spheres** into different simplicial complexes (see figure and table).
- TAE superior to other filtration functions in terms of separating the manifolds in this setting.

	TAE	PCA	t-SNE	SVD	Kernel density	Eccentricity
Best	**1.000**	1.217	1.060	1.227	6.609	4.333
Worst	**1.028**	2.920	1.181	2.920	8.583	6.333

Best average number of spheres in each node. Lower is better, 1 is optimal.



Mapper graphs for the best results in table (see left). Node color indicate average label value of which sphere that points in the node belong to – yellow is the large sphere, others are the smaller.

Discussion & Future work

- Initial results show the potential of TAEs for the Mapper.
- Currently investigating higher dimensional manifolds to confirm hypothesis.
- Creating new validation metrics to validate the graphs in terms of:
  - Manifold separation.
  - Point cloud connectivity.

Acknowledgements

This work was partially supported by Wallenberg Autonomous Systems Program (WASP).

[1] Michael Moor, Max Horn, Bastian Rieck, and Karsten Borgwardt. Topological autoencoders. arXiv preprint arXiv:1906.00722, 2019.

Topological Data Analysis and Beyond Workshop at the 34th Conference on Neural Information Processing Systems (NeurIPS 2020), Vancouver, Canada.