

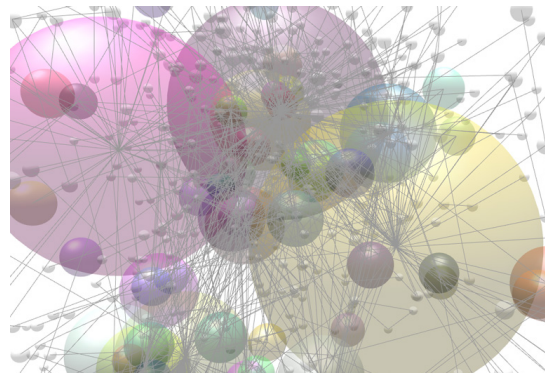


BA/MA:

High Dimensional Clustering

Clustering high dimensional datasets is a fundamental problem in computational geometry with significant impact on data science, big data and machine learning. Clustering is NP-hard even in low dimensions, but there are efficient approximation algorithms. However, in high dimensional datasets another problem occurs known as the “curse of dimensionality”.

In this thesis, you will explore the vast area of clustering high dimensional datasets. Specifically, you will investigate the connection between clustering, greedy permutations and r -nets. Aspiring to overcome the “curse of dimensionality” and provide an efficient clustering algorithm will be your main focus. In this journey you will encounter state-of-the-art techniques that may involve probabilistic polynomials, double embeddings etc.



Requirements: Basic knowledge of probability theory and algorithmic design. Mathematical maturity will be an advantage.

Interested? Please contact us for more details!

Contacts

- Zeta Avarikioti: zetavar@ethz.ch, ETZ G95
- Yuyi Wang: yuwang@ethz.ch, ETZ G94