Geometric Data Analysis

0. Overview and Introduction

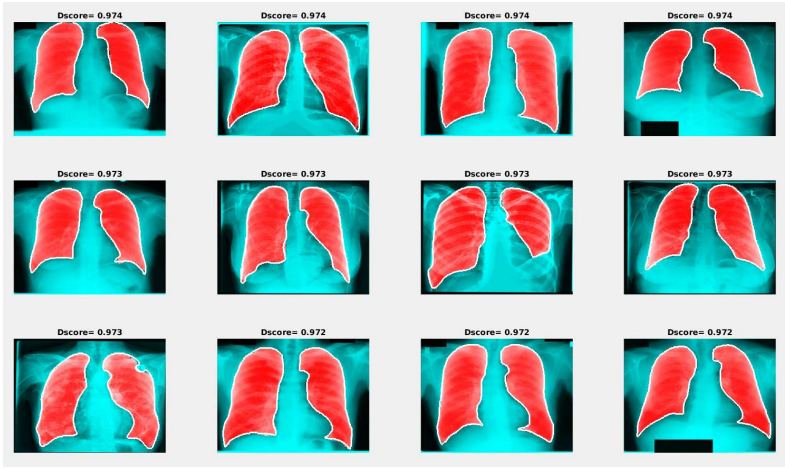
Ioannis Emiris

Dept Informatics & Telecoms, National Kapodistrian U. Athens
ATHENA Research & Innovation Center, Greece

Fall 2020

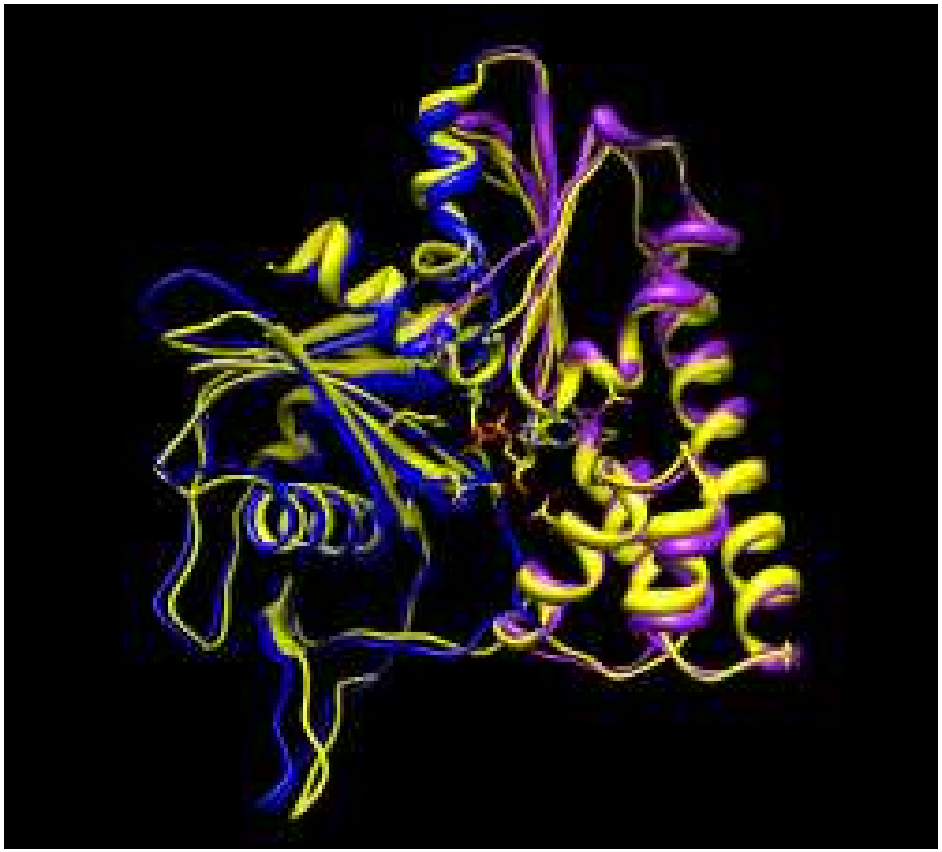
Machine vs humans

What can the computer do better than humans?



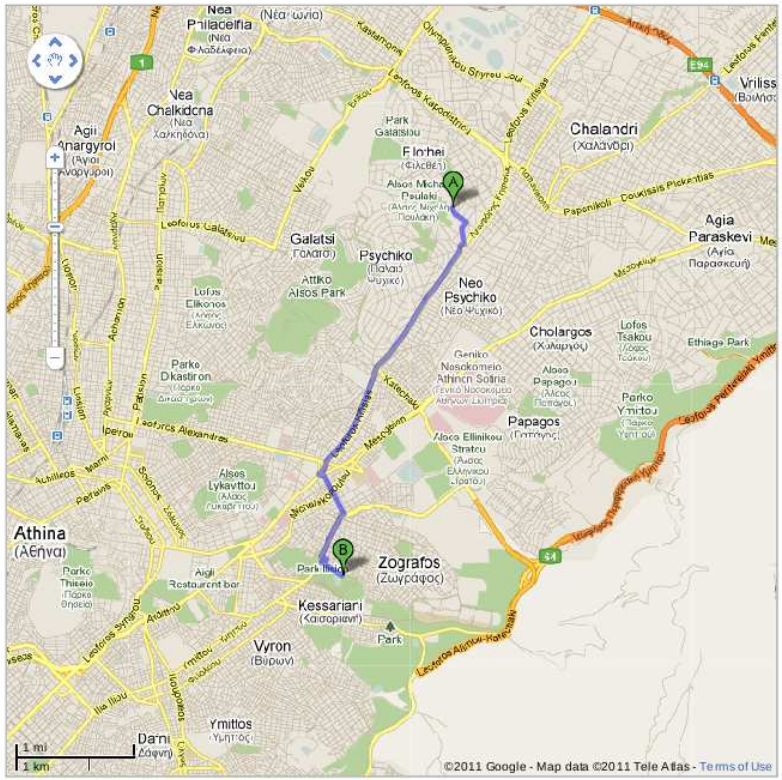
Data crunching. . . also some tasks in modern geomertic computing. . .

Mining complex objects

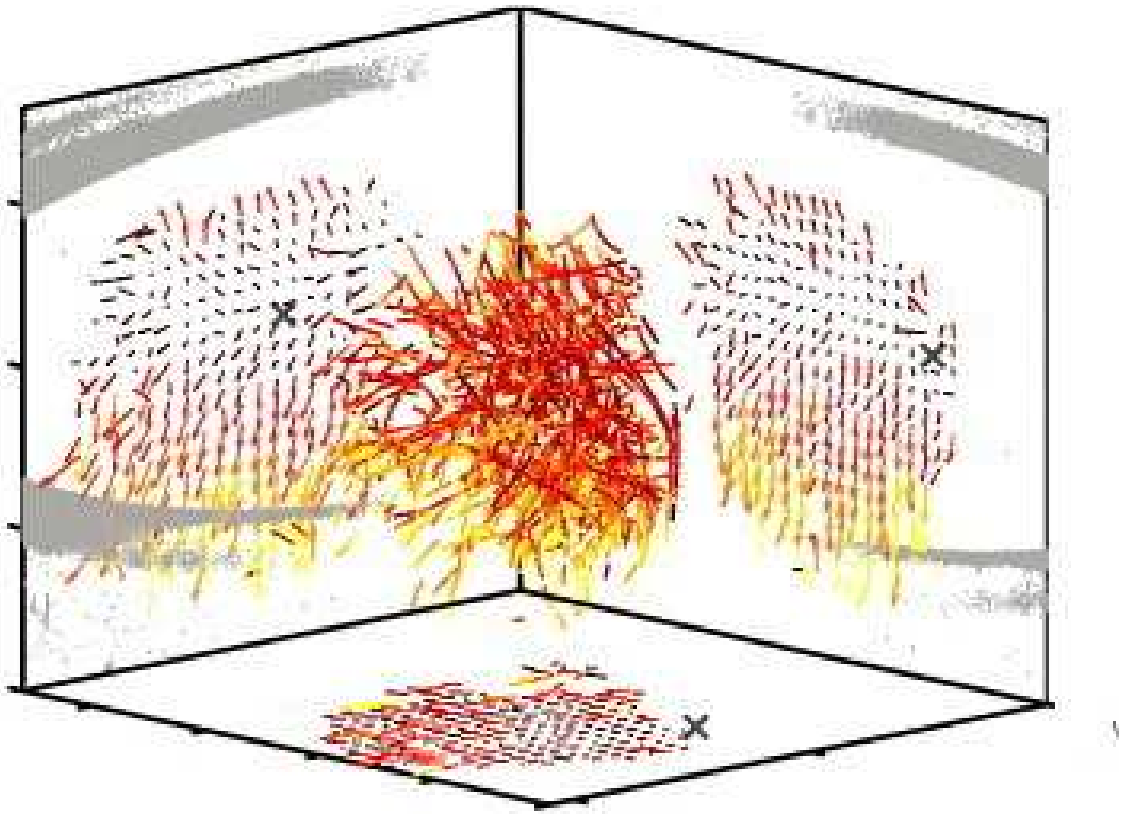


Massive (geographic) data

Stefanou Delta/Στέφανου Δέλτα το Πανεπιστημι... http://maps.google.com/maps?f=d&source=s_d&s...

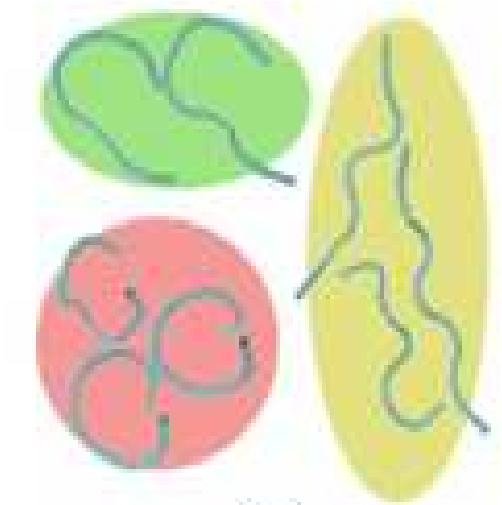


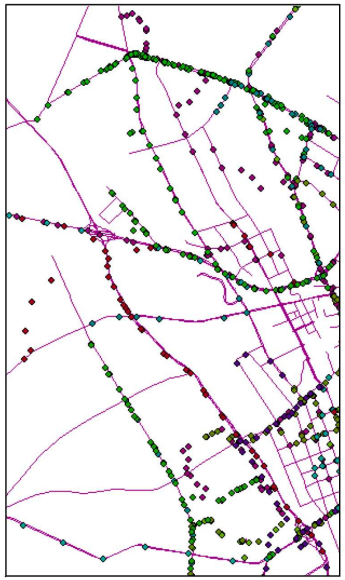
Curse of dimensionality



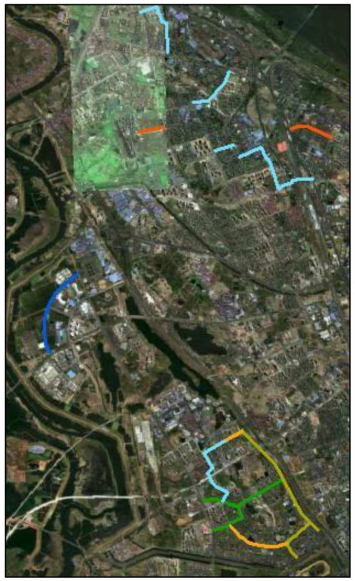
Data management and clustering

One-dimensional objects



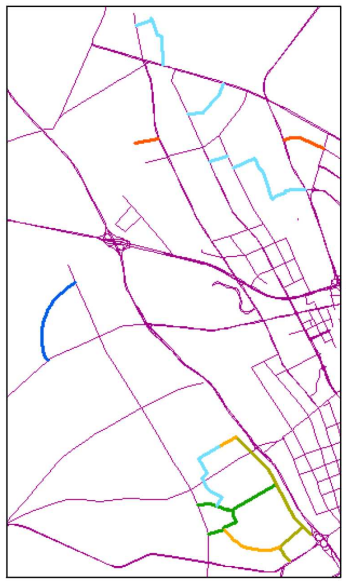


(b)



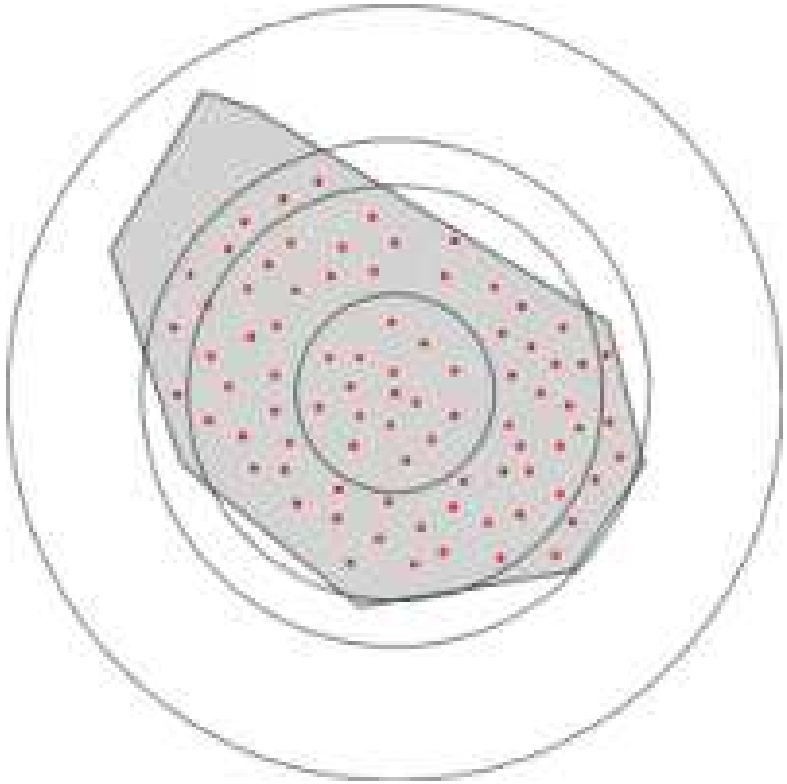
(d)

(a)

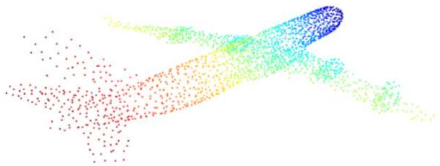


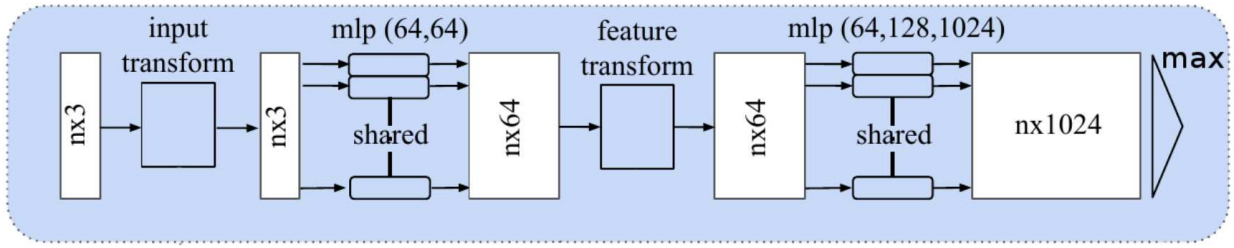
(c)

Sampling



Representation / Encoding

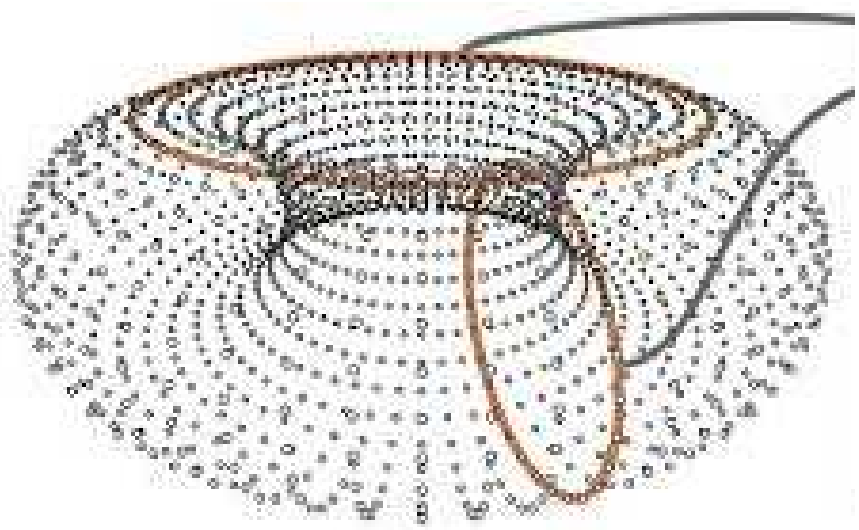


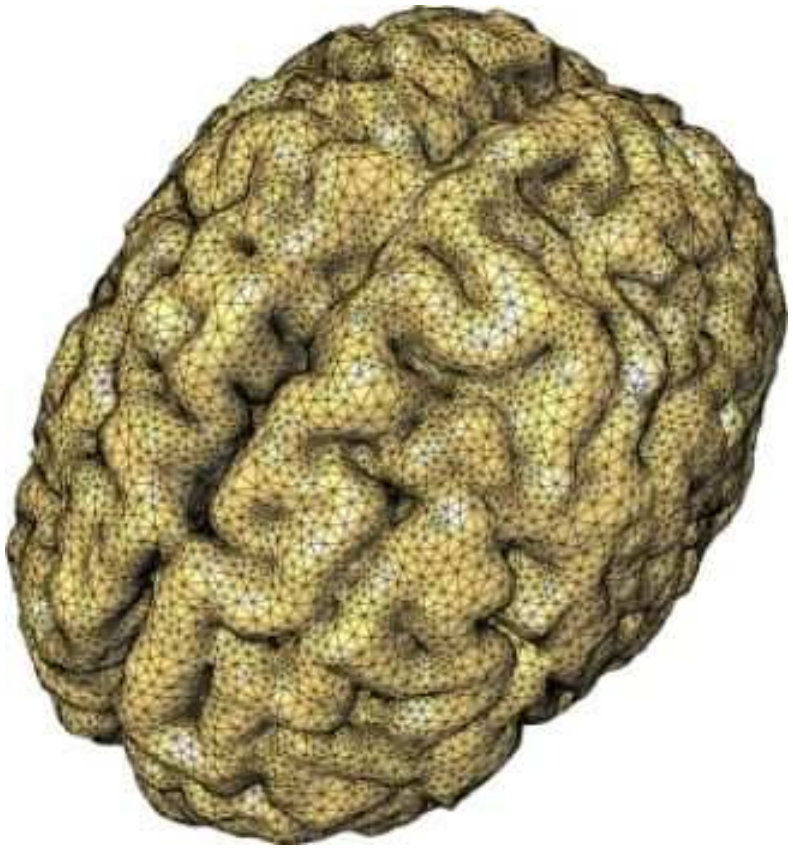


Not covered

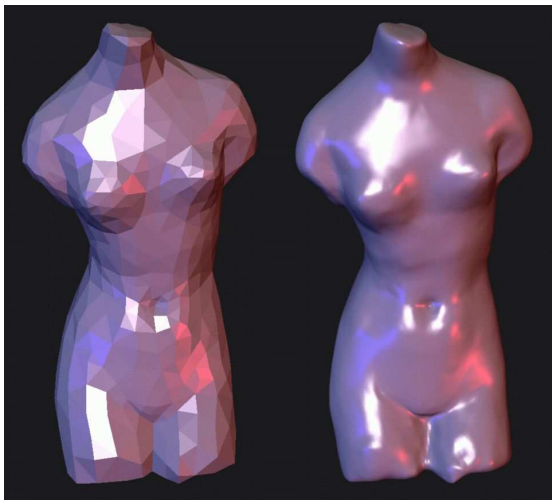
Several topics not covered, for example:

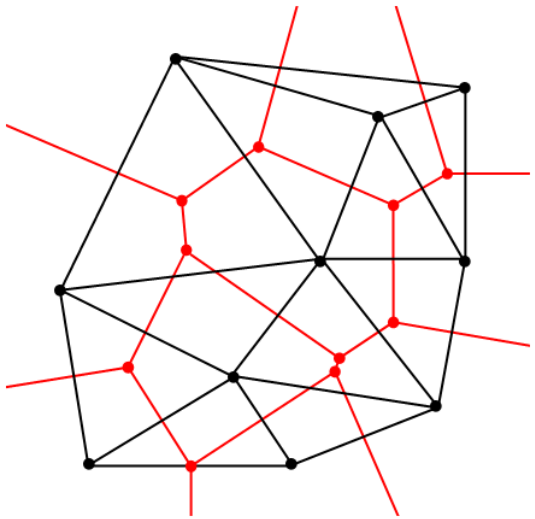
Topology of data, and Visualization





Graphics, movies





Course schedule

13 3-hour classes: Overview:
– Polytopes, Sampling, Random walks, Volume, optimization. Applications to finance [Emiris, Chalkis]
– Search, mining, clustering: LSH, Randomized Projections, JL lemma, Dimensionality reduction, Clustering [Emiris, Psarros]
– Graph Neural networks, Representation learning: node and Graph classification, Sets embeddings, Data challenge [Vazirgiannis]

Books on the web:
- Leskovec,Rajaraman,Ullman: Mining Massive Datasets: `www.mmds.org`
- Blum, Hopcroft, Kannan: Foundations of Data Science

Grade: 3 Course presentations or Projects (solo or pairs).
Projects: paper implementation or programming project.