



UNIVERSITÀ
degli STUDI
di CATANIA

DIPARTIMENTO
di ECONOMIA
e IMPRESA

SEMINARIO

Eugene Demidenko

*Department of Biomedical Data Science,
Geisel School of Medicine at Dartmouth
Department of Mathematics, Dartmouth College*

The next-generation K-means algorithm

Abstract: The goal of the talk is to discuss some extensions of the classic K-means algorithms to situations when underpinning assumptions, such equal variance /homogeneity or independence within cluster, breakdown. A statistical model-based approach is the basis for our developments. The following problems are discussed: (1) computation of the p-value against false positive clustering, (2) broken-line algorithm for estimation of the number of clusters, (3) robust clustering via Laplace distribution, (4) semi-parametric clustering norm, (4) cluster-wise regression, (5) clustering of dependent observations (heterogeneous data). Two applications illustrate our developments: (1) classification of Atomic Force Microscopy (AFM) images for detection of bladder cancer cells found in urine samples ,and (2) three-dimensional sphere gene network dynamic visualization of human DNA.

Giovedì 29 Settembre 2022, ore 15.30
Aula 5 - Palazzo delle Scienze, Corso Italia n°55, Catania
