Workshop on Algorithms for Modern Massive Data Sets

Stanford University and Yahoo! Research

June 21–24, 2006

Objectives: Explore novel techniques for modeling and analyzing massive, high-dimensional, and nonlinear-structured data. Bring together computer scientists, computational and applied mathematicians, statisticians, and practitioners to promote cross-fertilization of ideas

Theory: large scale numerical linear algebra; kernel-based nonlinear structure extraction; tensor-based multilinear structure extraction; geometrical and topological techniques; missing value estimation; sampling-based algorithms

Applications: analyzing microarray data and high-throughput chemical data in pharmaceutical applications; identifying gene products, elucidating protein folding pathways; detecting and classifying cancer; modeling combinatorial structure of large social, computer, and communication networks; identifying potential terrorist cells in communications networks; identifying noisy images of targets and faces in realistic settings; improving internet search engines; analyzing remote sensing data for environmental planning, weather forecasting, and public health contamination

Organizers: Gene Golub, Michael Mahoney, Petros Drineas, Lek-Heng Lim

Sponsors: National Science Foundation, Stanford Computer Forum, Yahoo! Research

http://mmds.stanford.edu