

A Geometric Perspective on Learning Theory and Algorithms

Partha Niyogi

The University of Chicago, Hyde Park, Chicago,IL 60637.

Increasingly, we face learning problems in very high dimensional spaces. We proceed with the intuition that although natural data lives in very high dimensions, they have relatively few degrees of freedom. One way to formalize this intuition is to model the data as lying on or near a low-dimensional manifold embedded in the high dimensional space. This point of view leads to a new class of learning algorithms that are “manifold motivated” and a new set of theoretical questions that surround their analysis. A central construction in these algorithms is a graph or simplicial complex that is data-derived and we will relate the geometry of these to the geometry of the underlying manifold. Applications to embedding, clustering, classification, and semi-supervised learning will be considered.