A new perspective on machine learning

A central problem of machine learning is the following. Given data of the form $\{(x_i, y_i) = z_i\}$ drawn from an unknown probability distribution μ, find a functional relationship between x's and y's. Alternatively, denoting $\mathbb{E}_\mu(y|x)$ by $f(x)$, approximate f given samples of the form $f(x_i) + \epsilon_i$, where x_i's are drawn randomly from an unknown (marginal) distribution μ^* and ϵ_i are random noise variables from another unknown distribution. Most of the approximation theory results deal with approximation of functions on known domains, requiring that the data be asymptotically dense on this domain. This assumption is usually not satisfied on real data. Therefore, traditional machine learning has focused on various elaborate algorithms which, however, are often ad hoc, and do not have an obvious direct connection with the problem of function approximation. We discuss some ideas motivated by diffusion geometry to address the central problem more directly. In particular, the same tools useful in function approximation can be used also for the estimation of the probability distributions μ or μ^* in a dual manner.

H. N. Mhaskar
Claremont Graduate University,
Claremont, CA 91711 U.S.A.