

Adaptive Neighborhood Definitions for Local Learning
Maya R. Gupta (presenting author), Eric K. Garcia and Erika Chin
University of Washington
Seattle, WA 98195
gupta@ee.washington.edu
idl.ee.washington.edu

Local learning methods, such as local linear regression and nearest neighbor classifiers, form predictions based on a local neighborhood of a test point. The most common neighborhood is a fixed number of nearest training samples to the test sample. In this research, we propose instead adapting the neighborhood size to the local spatial distribution of the data. The proposed adaptive neighborhood definitions do not require cross-validation. The term *enclosing neighborhood* is introduced to describe a set of neighbors whose convex hull contains the test point if possible. We prove that enclosing neighborhoods yield bounded estimation variance under some assumptions. Three such enclosing neighborhood definitions are presented and contrasted: Sibson's *natural neighbors* [1], a proposed variant we call *natural neighbors inclusive*, and a new neighborhood definition with the descriptive name *enclosing k-NN*. The enclosing k-NN neighborhood corresponds to the smallest k number of neighbors needed to enclose the test sample. The effectiveness of these neighborhood definitions is tested with local linear regression and regularized local linear regression for estimating LUTs for color management of printers and for custom color enhancements. The proposed neighborhood methods reduce color management error by 10-50% compared with the standard approach. An asymptotic analysis establishes the neighborhood size in the limit as the number of training samples $n \rightarrow \infty$.

References

[1] **R. Sibson** "A brief description of natural neighbour interpolation," appeared in *Interpreting Multivariate Data*, John Wiley, 1981.

Topic: learning algorithms

Preference: oral