

Optimal weighted nearest neighbor classifiers

DR. RICHARD SAMWORTH

Cambridge University

Classifiers based on nearest neighbors are perhaps the simplest and most intuitively appealing of all nonparametric classifiers. Arguably the most obvious defect with the k -nearest neighbor classifier is that it places equal weight on the class labels of each of the k nearest neighbors to the point being classified. Intuitively, one would expect improvements in terms of the misclassification rate to be possible by putting decreasing weights on the class labels of the successively more distant neighbors. In this talk, we determine the optimal weighting scheme, and quantify the benefits attainable. Notably, the improvements depend only on the dimension of the data, not on the underlying population densities.