A GENERALIZATION OF THE k-NEAREST NEIGHBOR RULE

E. A. Patrick and F. P. Fischer, II
School of Electrical Engineering
Purdue University, Lafayette, Indiana

ABSTRACT

A family of supervised, nonparametric decision rules, based on tolerance regions, is described which includes the k-Nearest Neighbor decision rules when there are two classes. There are two practical reasons for doing so: first, a family of decision rules similar to the k-Nearest Neighbor rules can be specified which applies to a broader collection of pattern recognition problems. This is because, in the general class of rules, constraints are weakened between the number of training samples required in each training sample set and the respective a priori class probabilities; and a discrete loss function weighting the importance of the finite number of ways to make a decision error can be introduced.

Second, within the family of decision rules based on tolerance regions, there are decision rules which have a property allowing for preprocessing of the training set data resulting in significant data reduction.

Theoretical performance for a special case is presented.