Nearest Neighbors

In $k-NN$ classification, the output is a class membership. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors (k is a positive integer, typically small). If $k = 1$, then the object is simply assigned to the class of that single nearest neighbor.

Properties

- $k-NN$ is a special case of a variable-bandwidth, kernel density "balloon" estimator with a uniform kernel.

- For multi-class $k-NN$ classification, Cover and Hart (1967) prove an upper bound error rate of

$$R^* \leq R_{kNN} \leq R^* \left(2 - \frac{MR^*}{M-1} \right)$$

where R^* is the Bayes error rate (which is the minimal error rate possible), R_{kNN} is the $k-NN$ error rate, and M is the number of classes in the problem. For $M = 2$ and as the Bayesian error rate R^* approaches zero, this limit reduces to "not more than twice the Bayesian error rate".