Nearest Neighbor Density Estimation

The kernel density estimators are easy to use.

However they are computationally expensive.

Consider 2-class problem with n1 + n2 = n training samples

If we use Gaussian window function at any x we need to compute n Gaussians

If we can model both class conditional densities as Gaussian, the required computation is much less.

Another issue is the size of the volume element.

Choosing the value for h is difficult

Sometimes one may choose different h in different parts of feature space.

Kernel density estimators are the most popular nonparametric estimator.

A different approach to non-parametric density estimation is the k-nearest neighbor approach.

Here we do not have to choose the size of h.

Instead we choose k and find V to enclose the k-nearest neighbor of x

Then we take

The nearest neighbor density estimation is closely related to nearest neighbor classifier.

Consider a 2-class problem with Prior Probabilities and class conditional densities , i =0, 1

Let be the overall density of feature vector

Suppose there are n data samples with being from class-I, i =0, 1

We do k-nearest neighbor estimation of f. Suppose the needed volume is V.

Suppose in this volume, there are samples of class-i, i =0, 1

Now using the same volume element, we estimate densities f as well as , i =0,1

, i= 0, 1 and

The estimates for prior probabilities would be

Using the estimates, compute the posterior probabilities

Now if we want to implement Bayes classifier x would be put in class j if

which implies

Thus the Bayes classifier with these estimated densities is the k-nearest neighbor classifier.

NN Classifier :

The error rate will never be more than twice of Bayes’ error rate

Bayes error is then the error rate of NN classifier 2

Asymptotically as the number of sample goes to infinity the NN classifier error rate in worst case is bounded 2\*Bayes’s error rate.