



1. Using the reference data set identify the  $k$  nearest neighbors in  $Y$  variable space for each observation.
2. Derive (normalized) centroid in  $X$  variable space based on the  $k$  nearest neighbors in  $Y$  variable space, for each observation.
3. For each target observation calculate the distance to each nearest neighbor centroid after normalizing the target  $X$  variable set with respect to each of the observations in the reference dataset.