CS497, Fall 07

Assignment for Shai Ben-David's lectures

Due November 29th, 2007 at 4:30pm

We consider the axiomatic approach to clustering. Recall that

- A distance function over some domain X is a mapping, $d : X \rightarrow \Re^+$ such that, for all $x, y \in X$, $d(x, y) = d(y, x)$, and $d(x, y) = 0$ *iff* $x = y$.
- A clustering function for some domain set X is a mapping from the set of distance functions over X to the set of partitions of X.
- **Clustering Axioms:**
  1. A clustering function, F, satisfies *Scale Invariance (SI)* if, for every distance functions, d, d', if there exist a positive real number λ, such that for every $x, y \in X$, $d'(x, y) = \lambda d(x, y)$, then $F(d) = F(d')$ (that is, F outputs the same clustering of X when it gets as input the distance function d and when its input is d').
  2. Given a distance function, d, and a clustering (i.e. a partition) of X, $C = (C_1, \ldots, C_k)$ (where the C_i's are disjoint subsets of X and their union equals X), a distance function, d', is a *C-locally consistent transformation* of d if there exist positive real numbers, $\lambda_0 \geq 1$ and $\lambda_1, \ldots, \lambda_k, \leq 1$ such that for every $1 \leq i \leq k$ and every x, y, that are members of C_i, $d'(x, y) = \lambda_i d(x, y)$ and for every x, y that belong to different C_i's, $d'(x, y) = \lambda_0 d(x, y)$. We say that F satisfies *Local Consistency (LC)* if for every d and every d' that is an $F(d)$- locally consistent transformation of d, $F(d) = F(d')$.
  3. F satisfies *Richness* if for every partition C of X there exists a distance function d such that $F(d) = C$.
  4. For a natural number. k, F satisfies *k-Richness (k-R)* if for every partition C of X into k subsets (that is, $C = (C_1, \ldots C_k)$), there exists a distance function d such that $F(d) = C$.

Consider clusterings into 2 clusters and the axioms: SI, LC and 2-R.

1. Define a clustering function that satisfies all of these three axioms.
2. For each pair of these axioms define a clustering function that satisfies these two axioms but not the third.