




Artificial Intelligence Programming

Instance Based Learning

Chris Brooks

Department of Computer Science
University of San Francisco



Instance-Based Learning

- So far, all of the learning algorithms we've studied construct an explicit hypothesis about the data set.
- This is nice because it lets us do a lot of the training ahead of time.
- It has the weakness that we must then use the same hypothesis for each element in the test set.
- One way to get around this is to construct different hypotheses for each test example.
 - Potentially better results, but more computation needed at evaluation time.
- We can use this in either a supervised or unsupervised setting.

k-nearest neighbor

- The most basic instance-based method is k-nearest neighbor.
- Assume:
 - Each individual can be represented as an N-dimensional vector: $\langle v_1, v_2, \dots, v_n \rangle$.
 - We have a distance metric that tells us how far apart two individuals are.
 - Euclidean distance is common:
$$d(x_1, x_2) = \sqrt{\sum (x_1[i] - x_2[i])^2}$$