Wednesday, September 15, 2021

Cs506 lec3

Data representation

- Vectors
- Graphs edges' connections are represented by adjacency matrixes (1 when there is an edge 0 when there is't)
- Adjacency list more efficient 1:{2,3} 1 is connected with 2 and 3
- Text
- Strings list of characters

Types of learning

Supervised learning and unsupervised learning

Similarity and distance
Feature space of all possible values for the set of features in our data.

Distance:
Compare the data points given
Dissimilarity function returns a small value they are similar if output is large then dissimilar.

Distance
D is a distance function if and only if :
d(i,j) if and only if I=j
d(i,j) = d(j,i)
d(i,j) <= d(i,k) +d(k,j) if going through another point can only add distance to the path

It's intuitive to understand and it makes sense.

Minkowski distance
When 0<p<1 it's no longer a distance function