

Monday September 23rd, 2013

Last time: Intro to Fellegi-Sunter

Today: Fellegi-Sunter; Transitivity; Single Linkage

Announcements:

- HW 3 due Friday 9/27
- Start thinking about project ideas

Fellegi-Sunter Review:

Using weights to calculate the Ratio R :

Expectation-Maximization Algorithm:

Starting values? What could we use?

What do we do with these probabilities once we have them?

Transitivity:

How would we code this?

How is this a clustering problem?

Hierarchical (Agglomerative) Clustering: Single Linkage, in particular

```
dist.obj<-1-as.dist(prob.matrix)
hc<-hclust(dist.obj,method="single")
plclust(hc)
groups<-cutree(hc,h=xxx)  ##can also do k = xxx for a specific \# o groups
```