labs: project updates (5pts)

attendance is mandatory

• individual contributions
• previous goal
• specificity and scope of next goal
progress report due
1:00 am Monday, 11/16

this is a hard deadline:
no late days allowed!

progress presentations moved to Nov 23
(3-5 slides)
intro to recommender systems
Recommender Systems

- collaborative filtering
- content-based filtering

amazon

netflix

pandora
preliminaries
Clustering
Clustering
Hierarchical Clustering
Hierarchical Clustering
Hierarchical Clustering
Hierarchical Clustering
Hierarchical Clustering
Hierarchical Clustering
Hierarchical Clustering
Hierarchical Clustering
Hierarchical Clustering
Hierarchical Clustering
Hierarchical Clustering
Hierarchical Clustering
Hierarchical Clustering
Hierarchical Clustering
k-means Clustering

- pick k initial points (*means*)
- for every remaining point x
  - assign x to the cluster with the closest mean
- recompute means
k-means Clustering
k-means Clustering
measuring similarity and distance
Jaccard Similarity

\[
\text{sim}(S, T) = \frac{|S \cap T|}{|S \cup T|}
\]

for given sets \( S \) and \( T \)
Jaccard Similarity
Distance Measures

distance \( d(x,y) \) satisfies

- \( d(x,y) \geq 0 \) with equality exactly when \( x = y \)
- \( d(x,y) = d(y,x) \)
- triangle inequality \( d(x,y) \leq d(x,z) + d(z,y) \)

**example:**

- **Jaccard Distance** = 1 - Jaccard Similarity
- \( d(X,Y) = 1 - \text{sim}(X, Y) \)
  
  \[ = 1 - \frac{|X \cap Y|}{|X \cup Y|} \]
Cosine Distance

- treat points as “directions”
- distance is the angle between them (0-180)
Cosine Distance

distance $d(x,y)$ satisfies

- $d(x,y) \geq 0$ with equality exactly when $x = y$
- $d(x,y) = d(y,x)$
- triangle inequality $d(x,y) \leq d(x,z) + d(z,y)$

$$\theta \leq \theta_1 + \theta_2$$
Cosine Similarity

- treat points as “directions”
- similarity is $\cos \theta$ and lies between -1 and 1
User-user CF

idea: find users with high agreement on rated items

\[ p_{u,i} = \bar{r}_u + \sum_{u' \in N} \frac{s(u, u')(r_{u',i} - \bar{r}_{u'})}{\sum_{u' \in N} \left| s(u, u') \right|} \]