

Carnegie Mellon University

University Undergraduate Course: Data Mining

<https://www.stat.cmu.edu/~ryantibs/datamining/>
(retrieved 11 March 2024)

Data Mining: Spring 2013

Statistics 36-462/36-662

Instructor: Ryan Tibshirani (ryantibs at cmu dot edu)

Teaching assistants:
Li Liu (liu1 at andrew dot cmu dot edu)
Cong Lu (congl at andrew dot cmu dot edu)
Jack Rae (jwr at andrew dot cmu dot edu)
Michael Vespe (mvespe at andrew dot cmu dot edu)

Lecture times: Tuesdays and Thursdays 1:30-2:50pm, Porter Hall 125C

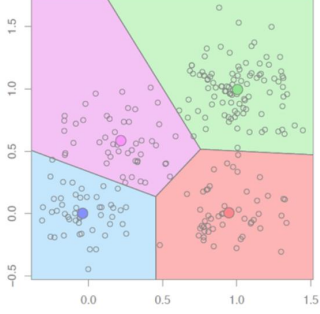
Recitation times: Wednesdays 5-6pm, Porter Hall 125C

Office hours: RT: Tuesdays 3-4pm, Baker 229B
LL: Wednesdays 4-5pm, FMS 320
CL: Fridays 11am-12pm, Wean 8110
JR: Wednesdays 11am-12pm, Wean 8110
MV: Mondays 5-6pm, FMS 320

Course syllabus: [PDF](#)

Go to: [Lectures](#) | [Recitations](#) | [Assignments](#) | [Schedule](#) | [Examples for extra credit](#)

Lecture notes



<https://www.stat.cmu.edu/~ryantibs/datamining/lectures/06-clus3-marked.pdf>
(retrieved 11 March 2024)

06-clus3-marked.pdf

<https://www.stat.cmu.edu/~ryantibs/datamining/lectures/06-clus3-marked.pdf>

1 of 32

Clustering 3: Hierarchical clustering (continued);
choosing the number of clusters

Ryan Tibshirani
Data Mining: 36-462/36-662

January 31 2013

06-clus3-marked.pdf

<https://www.stat.cmu.edu/~ryantibs/datamining/lectures/06-clus3-marked.pdf>

Minimax linkage

Minimax linkage² is a newcomer. First define radius of a group of points G around X_i as $r(X_i, G) = \max_{j \in G} d_{ij}$. Then:

$$d_{\text{minimax}}(G, H) = \min_{i \in G \cup H} r(X_i, G \cup H)$$

Example (dissimilarities d_{ij} are distances, groups marked by colors): minimax linkage score $d_{\text{minimax}}(G, H)$ is the **smallest radius** encompassing all points in G and H . The center X_c is

(this is circle obviously)

06-clus3-marked.pdf

<https://www.stat.cmu.edu/~ryantibs/datamining/lectures/06-clus3-marked.pdf>

Minimax linkage example

Same data as before. Cutting the tree at $h = 2.5$ gives clustering assignments marked by the colors

Minimax doesn't create inversions.

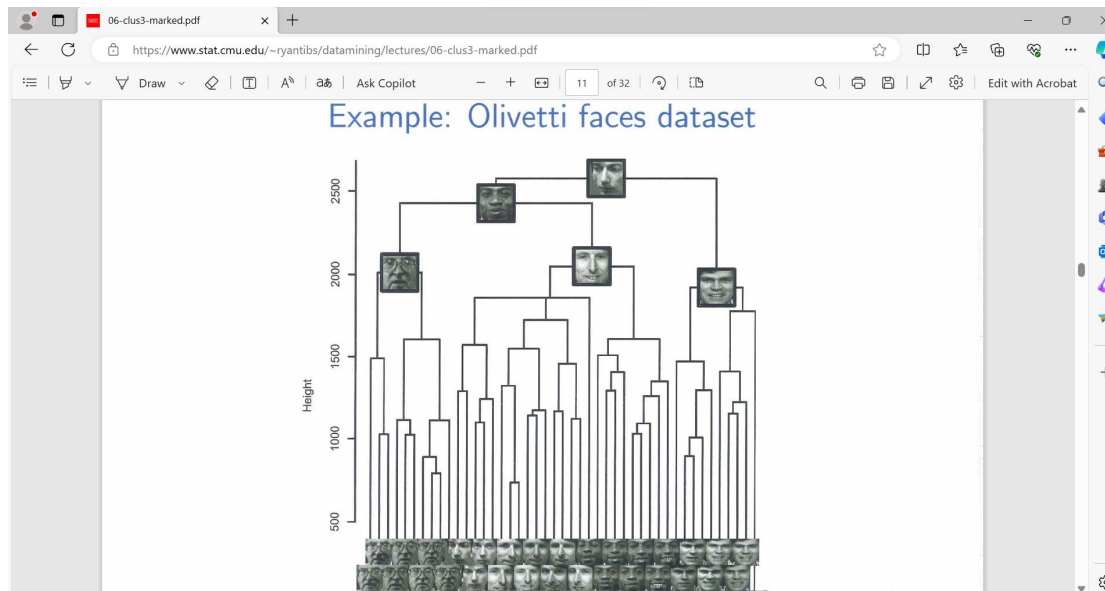
06-clus3-marked.pdf

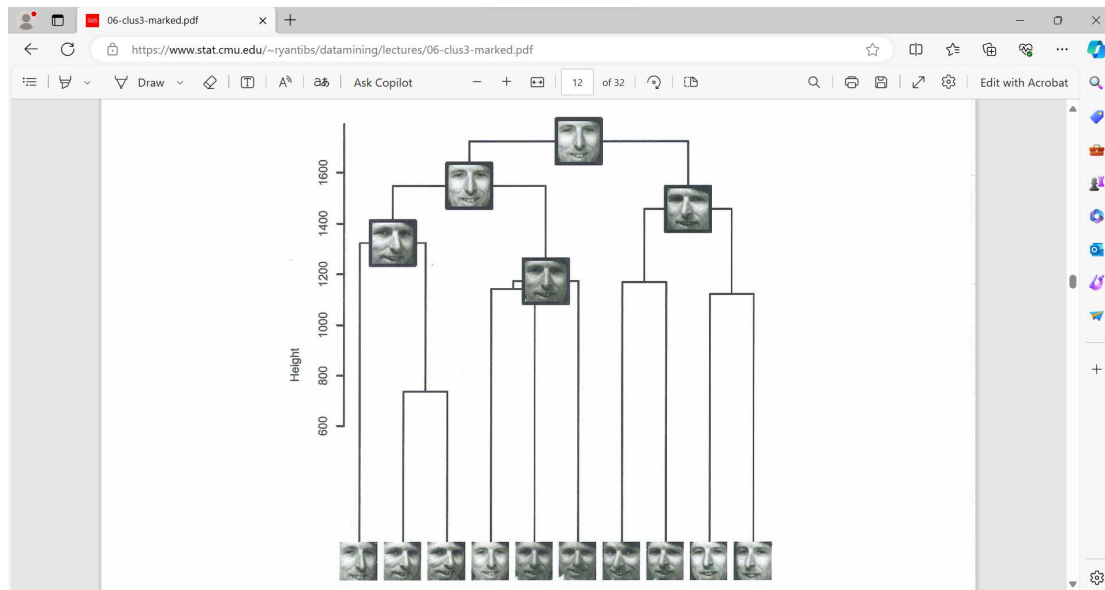
https://www.stat.cmu.edu/~ryantibs/datamining/lectures/06-clus3-marked.pdf

10 of 32

Properties of minimax linkage

- ▶ Cutting a minimax tree at a height h a **nice interpretation**: each point is $\leq h$ in dissimilarity to the center of its cluster. (This is related to a famous set cover problem)
- ▶ Produces dendograms with **no inversions**
- ▶ Unchanged by **monotone transformation** of dissimilarities d_{ij}
- ▶ Produces clusters whose **centers are chosen among the data points** themselves. Remember that, depending on the application, this can be a very important property. (Hence minimax clustering is the analogy to K -medoids in the world of hierarchical clustering)





Centroid and minimax linkage in R

The function `hclust` in the base package performs hierarchical agglomerative clustering with centroid linkage (as well as many other linkages)

E.g.,

```
d = dist(x)
tree.cent = hclust(d, method="centroid")
plot(tree.cent)
```

The function `protoclust` in the package `protoclust` implements hierarchical agglomerative clustering with minimax linkage

06-clus3-marked.pdf

https://www.stat.cmu.edu/~ryantibs/datamining/lectures/06-clus3-marked.pdf

14 of 32

Edit with Acrobat

Linkages summary

Linkage	No inversions?	Unchanged with monotone transformation?	Cut interpretation?	Notes
Single	✓	✓	✓	chaining
Complete	✓	✓	✓	crowding
Average	✓	✗	✗	
Centroid	✗	✗	✗	simple
Minimax	✓	✓	✓	centers are data points

Note: this doesn't tell us what "best linkage" is

What's missing here: a detailed empirical comparison of how they