

Hierarchical Clustering

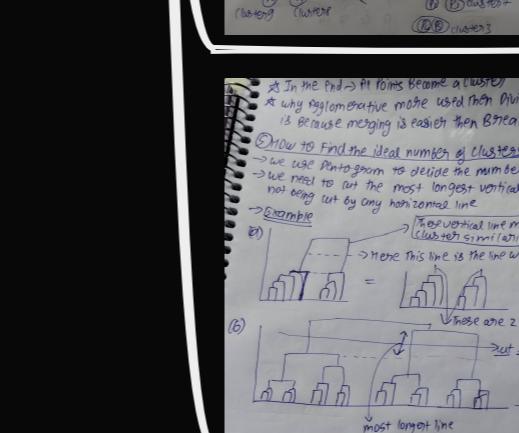
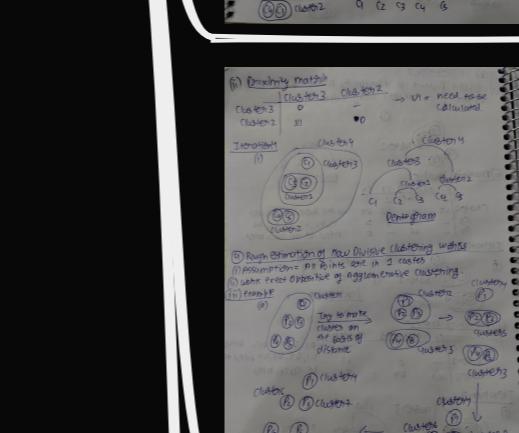
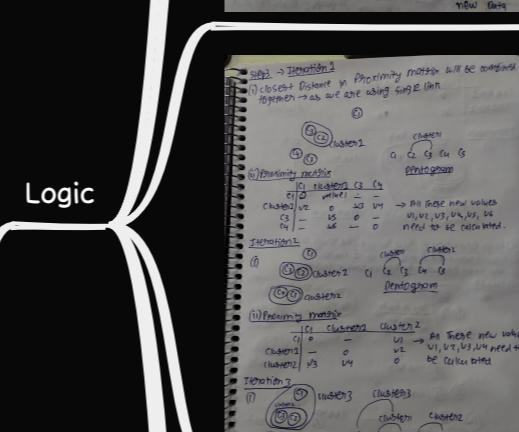
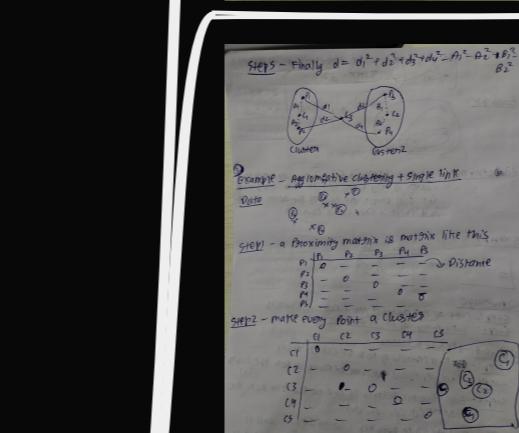
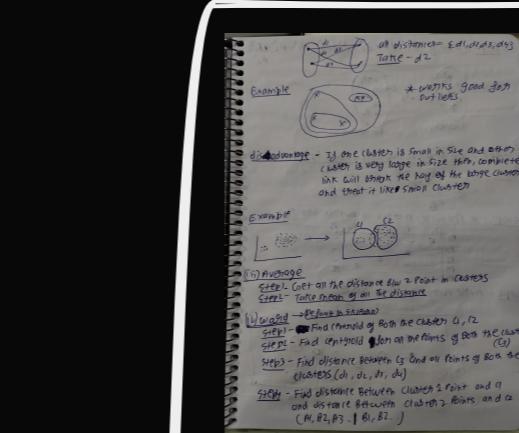
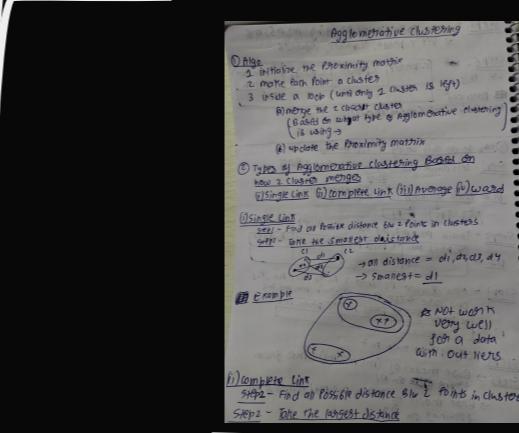
Agglomerative clustering mostly used

Types

Divisive clustering

video

github



important Hyper parameters

n_clusters

The number of clusters to find.

linkage

{'ward', 'complete', 'average', 'single'}, default='ward'

distance_threshold

float, default=None

The linkage distance threshold at or above which clusters will not be merged.

import scipy.cluster.hierarchy as shc

```
plt.figure(figsize=(10, 7))
plt.title("Customer Dendograms")
dend = shc.dendrogram(shc.linkage(data,
method='ward'))
```

Designing Dendograms

code

```
from sklearn.cluster import AgglomerativeClustering
```

```
cluster = AgglomerativeClustering(n_clusters=5, affinity='euclidean', linkage='ward')
labels_=cluster.fit_predict(data)
```

Benefit

widely applicable

disadvantage not applicable for large data set
(use proximity matrix)