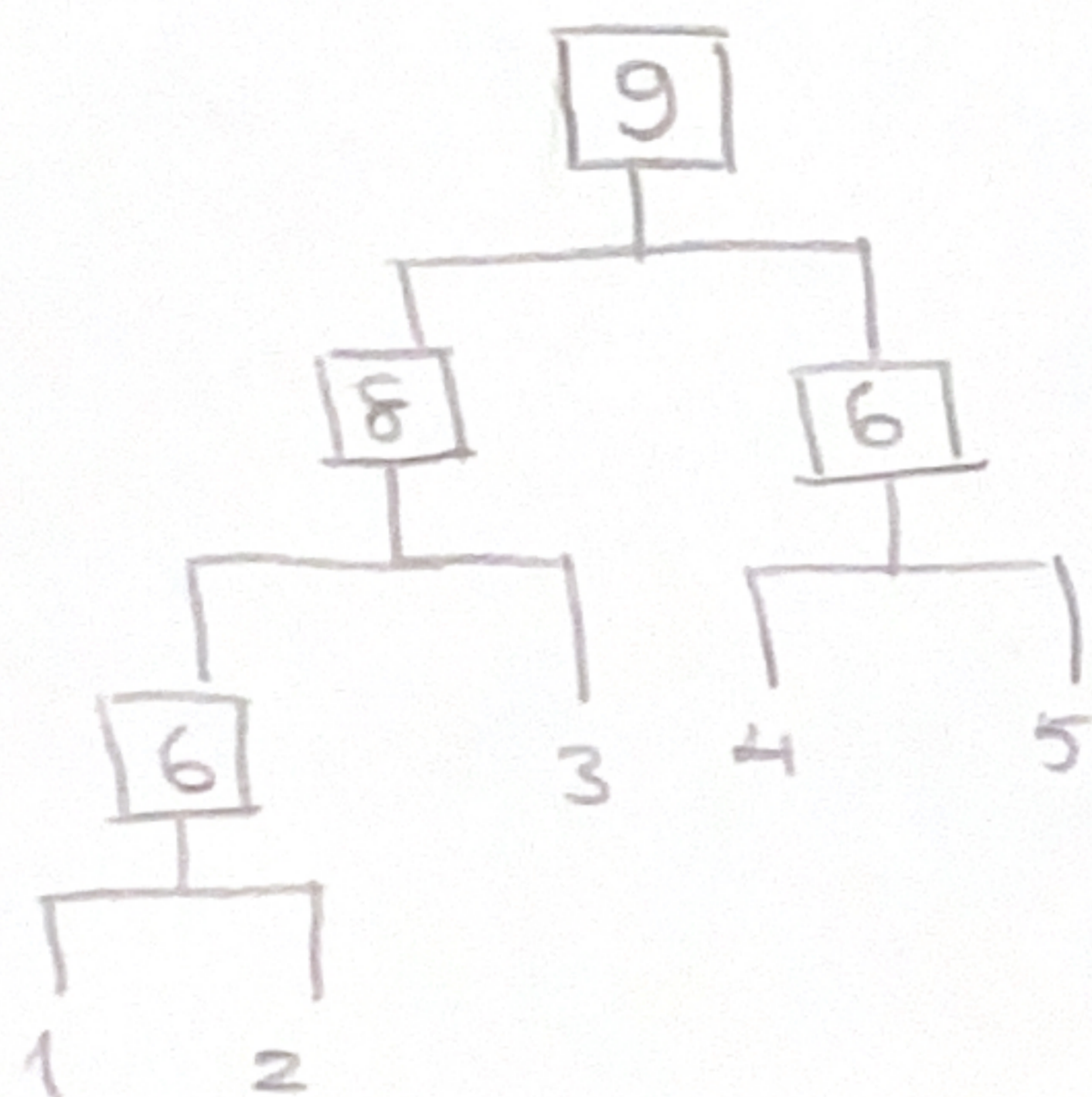


Unsupervised Learning

Hierarchical Clustering

Agglomerative (bottom up): Build the dendrogram (tree) from bottom level, merge most similar pair of clusters. Stop when all data points are merged into a single cluster.

Divisive (Top Down): Start with all data points in one cluster. Split it into child clusters. Divide them recursively. Stop when each cluster is only one data point.



Algorithm (Agglomerative) (D)

Make each data point D a cluster.

Compute all pairwise distances x_1, x_2, \dots, x_n

Repeat:

Find two clusters that are nearest

Merge them to new cluster c

Compute distance from c to all other clusters

Until

There's only one cluster

