

The background features a complex network of thin, intersecting lines in shades of brown and grey, creating a web-like structure. Scattered throughout are small, colored dots in green, blue, and orange. A prominent, larger, semi-transparent white shape, resembling a stylized letter 'A' or a cluster, is centered in the upper half of the image. On the left side, there is a small, rectangular inset showing a zoomed-in view of a cluster of orange and red dots.

Clustering Evaluation: Measuring Clustering Quality

Measuring Clustering Quality

- **Clustering Evaluation:** Evaluating the goodness of clustering results
 - No commonly recognized best suitable measure in practice
- **Three categorization of measures:** External, internal, and relative
 - **External:** Supervised, employ criteria not inherent to the dataset
 - Compare a clustering against prior or expert-specified knowledge (i.e., the ground truth) using certain clustering quality measure
 - **Internal:** Unsupervised, criteria derived from data itself
 - Evaluate the goodness of a clustering by considering how well the clusters are separated and how compact the clusters are, e.g., silhouette coefficient
 - **Relative:** Directly compare different clusterings, usually those obtained via different parameter settings for the same algorithm