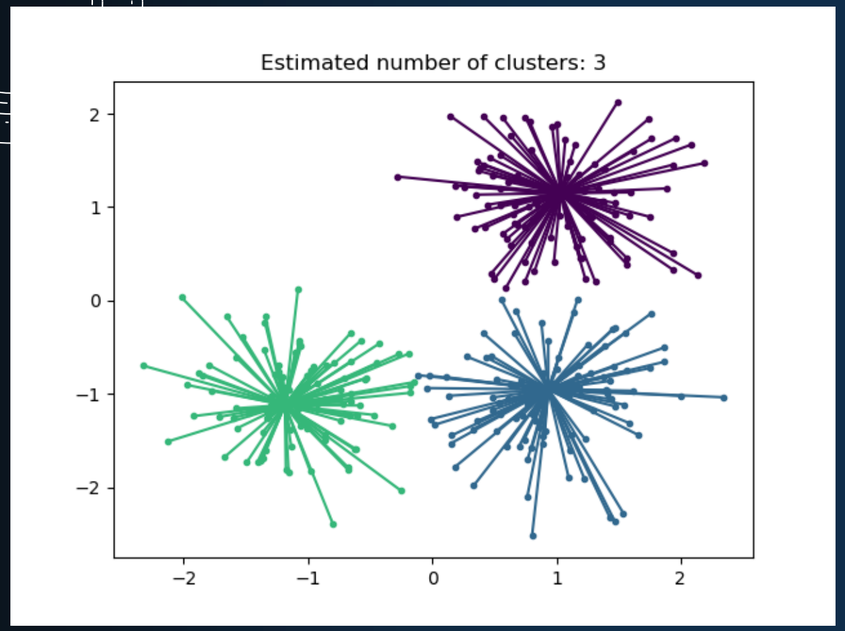
Affinity Propagation Clustering

MACHINE LEARNING CLUSTERING

Affinity Propagation Clustering

Introduction to Affinity Propagation

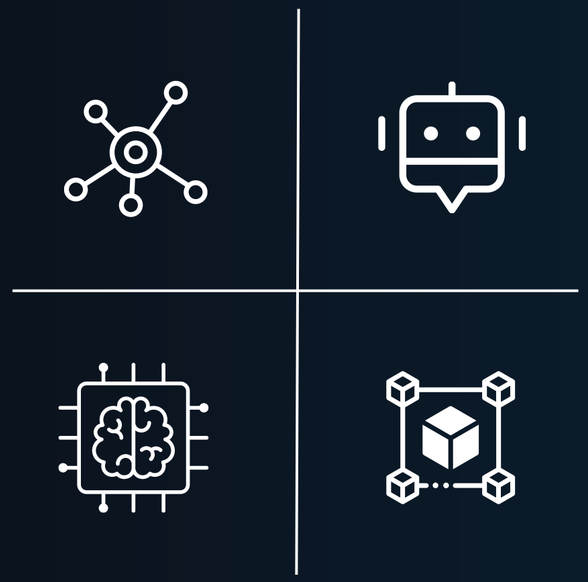
A clustering algorithm that identifies exemplars (representative data points) and forms clusters around them. Automatically determines the number of clusters based on data.



Key Concepts:

- **Exemplar:** A data point that represents a cluster.
- **Similarity Matrix (S):** Measures similarity between data points.
- **Responsibility (r(i, k)):** Suitability of point k as an exemplar for point i.
- **Availability (a(i, k)):** Appropriateness of choosing point k as an exemplar.

Mechanics of Affinity Propagation



1. Initialization:

Initialize responsibilities r(i, k) and availabilities a(i, k) to zero.

2. Update Responsibilities r(i, k).

3. Update Availabilities a(i, k).

4. Iteration:

Repeat responsibility and availability updates until convergence.

Application and Evaluation

Application:

- Compute similarity matrix from data.
- Run affinity propagation to identify exemplars and form clusters.

Advantages:

- Automatically determines the number of clusters.
- Can handle complex clustering structures.
- Suitable for large datasets.

Disadvantages:

- High memory and computational requirements.
- Sensitive to preference parameter.
- Potential convergence issues.
- Less scalable than simpler methods like K-means.
- Results can be difficult to interpret.

Thank You