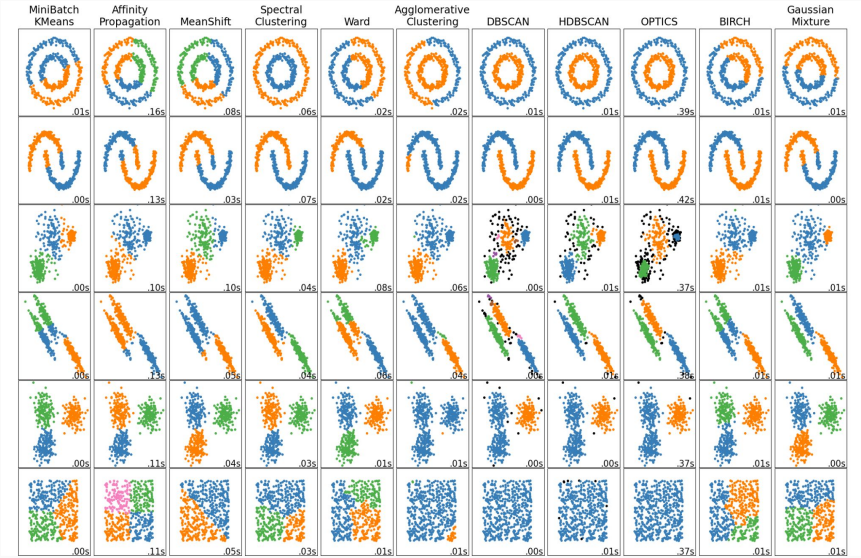
Affinity Propagation

Exemplar based clustering:

✔ Affinity Propagation is a message-passing algorithm where data points vote for the best representatives (exemplars).

✔ It views all data points as potential exemplars simultaneously.

✔ Unlike K-means, it does not require to specify the number of clusters(k) beforehand. It figures that out automatically.

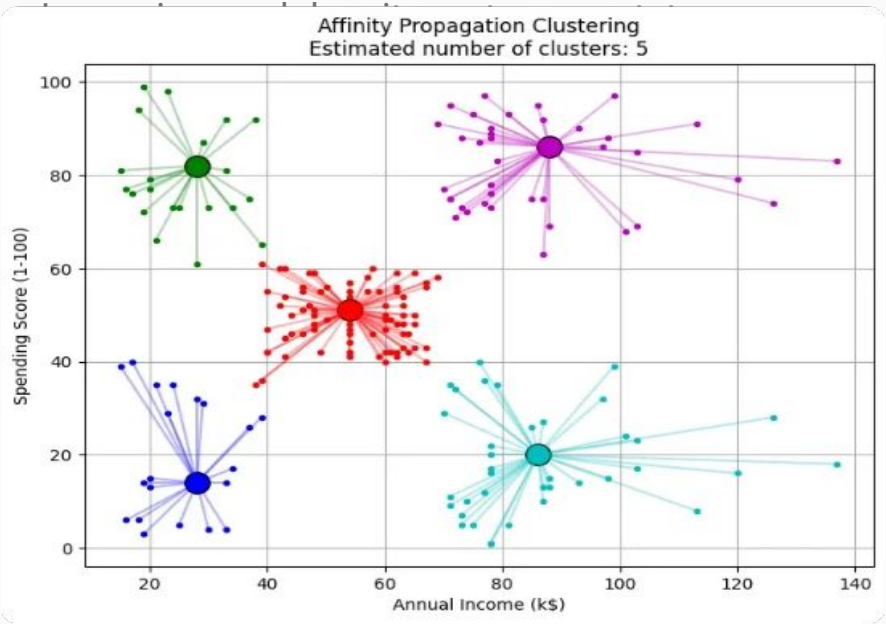


Message Passing:

The algorithm iterates by exchanging two types of real-valued messages between data points until convergence.

➜ **Responsibility r(i, k):**
Reflects how well-suited point k is to serve as the exemplar for point i, taking into account other potential exemplars.

➜ **Availability a(i, k):**
Reflects how appropriate it would be for point i to choose point k as its exemplar consider support from other points.



Advantages and Disadvantages:

Advantages:

1. No need to choose the number of clusters.

2. Finds real data point representatives (exemplars)

3. Can detect clusters of different shapes and sizes.

4. Works well when clusters have unequal sizes.

5. Good performance with moderate data sizes.

Disadvantages:

1. **Computational Cost:** Complexity is approximately O(N^2), making it slow for large datasets (N > 3000).

2. Not ideal for very high-dimensional data.

3. **Parameter Tuning:** The "Preference" and "Damping Factor" values can be non-intuitive to tune for specific results.

4. **Memory Intensive:** Requires storing dense similarity matrices.