Qeios

Peer Review

Review of: "Fixed-Sized Clusters k-Means"

Mohamed Abbas[1]

1. Alexandria University, Alexandria, Egypt

This paper presents a structured modification of k-means that enforces fixed cluster sizes using the Hungarian algorithm. While the method is mathematically sound, its practical validation is lacking. The study demonstrates only a single application (seating arrangement) without benchmarking against existing clustering methods or using standard evaluation metrics. Additionally, the $O(n^3)$ complexity limits its scalability, making it impractical for datasets beyond 5,000 points.

To strengthen the evaluation, the authors could benchmark their method using synthetic datasets with controlled cluster sizes (e.g., generated via make_blobs in sklearn). Comparisons with constrained k-means, spectral clustering, and DenMune on structured datasets, evaluated with NMI, ARI, or clustering accuracy, would provide deeper insights into its effectiveness.

For scalability, greedy assignment heuristics could be explored as a faster alternative to the Hungarian algorithm. A greedy approach ($O(n \log n) - O(n^2)$) could iteratively assign points to clusters while maintaining fixed sizes, significantly reducing computational cost while sacrificing some optimality. Despite these limitations, the approach fills a niche in constrained clustering, making it valuable for applications such as workforce scheduling, resource allocation, and balanced clustering in distributed processing.

Declarations

Potential competing interests: No potential competing interests to declare.

doi.org/10.32388/U5MAEW