ICRAMCS 2026

THE EIGHTH EDITION OF THE INTERNATIONAL CONFERENCE ON RESEARCH IN APPLIED MATHEMATICS AND COMPUTER SCIENCE
April 23-24-25, 2026 | Marrakech, Morocco

Optimal Cluster Estimation via Reference-Point Cross-validation Consensus Clustering

Communication Info

Authors:
Norin Rahayu SHAMSUDDIN[1]
Nor Idayu MAHAT [2]

[1] Integrated Simulation & Visualization Research Interest Group, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Kedah Branch, Malaysia
[2] Centre for Testing, Measurement and Appraisal, School of Quantitative Sciences, Universiti Utara Malaysia, Sintok, Malaysia

Keywords:
(1) Consensus clustering
(2) Reference point mechanism
(3) Optimal clusters estimation

Abstract

Estimating the optimal number of clusters, k remains challenging due to instability across resampling schemes and sensitivity to data structure. Although consensus clustering improves robustness [1,2], many ensemble approaches lack structured base-partition construction and principled k estimation [3,4]. A cross-validation–based consensus clustering framework incorporating a reference-point mechanism for stable k selection. Within each fold, 30% of observations are clustered across candidate k values, and internal validation indices :- Silhouette, GDI33, PBM, Point-Biserial, and Wemmert–Gançarski indices are used to determine fold-specific, $\check{k}_{\psi}$. Aggregation across folds yields stable partitions. Experiments on simulated and real datasets demonstrate improved robustness for unequal and moderately overlapping clusters while maintaining practical computational efficiency.

References

[1] Monti, S., Tamayo, P., Mesirov, J., & Golub, T. Consensus clustering: A resampling-based method for class discovery. *Machine Learning*, 52, 2003, 91–118.

[2] Strehl, A., & Ghosh, J. Cluster ensembles – A knowledge reuse framework for combining multiple partitions. *Journal of Machine Learning Research*, 3, 583–617, 2002.

[3] Vega-Pons, S., & Ruiz-Shulcloper, J*International Journal of Pattern Recognition and Artificial Intelligence*, 25(3), 337–372, 2011.

[4] Alqurashi, T., & Wang, W. Clustering ensemble method. *IEEE Access*, 8, 191146–191158, 2020.