Adaptive Neighborhood Search-Based Hybrid Clustering (ANSHC): A Robust Metaheuristic Framework for Partitional Data Clustering

Pratima Verma[1], Nipur Singh[2], Yugal Kumar[3] and Bal Ram[4]

{prtdreams@gmail.com[1], nipur@gkv.ac.in[2], yugalkumar.14@gmail.com[3], balrammlis@gmail.com[4]}

Department of Computer Science, Gurukula Kangri (Deemed to be) University, Haridwar, Uttarakhand, India[1]

Department of Computer Science, Gurukula Kangri (Deemed to be) University, Haridwar, Uttarakhand, India[2]

[3]School of Technology Management and Engineering, NMIMS, Chandigarh Campus, India[3]

Wadia Institute of Himalayan Geology Dehradun (An Autonomous Institute of DST, Govt. of India), India[4]

Abstract. Conventional clustering techniques like K-Means frequently encounter issues with local optima and are highly sensitive to initial conditions. This can result in less clustering outcomes, particularly when dealing with complex or high-dimensional datasets. This paper proposes a novel algorithm called Adaptive Neighborhood Search-based Hybrid Clustering (ANSHC) to address these limitations, which combines global meta-heuristic optimization with adaptive local search refinement. In the proposed ANSHC, an adaptive neighborhood search (ANS) is used so as to change the searching radius adaptively during the constrained optimization in a dynamic way between exploration and exploitation. Furthermore, we incorporate a hybrid local refinement operator, inspired by the K-Means centroid update procedure, to improve the precision of convergence and the compactness of clusters. Experimental studies on the ten benchmark datasets (Iris, Wine, Glass, WBC, Balance, and Thyroid) indicate that ANSHC consistently outperforms classical clustering algorithms and recent metaheuristic-based techniques PSO, GWO, and ChOA. The experimental result shows that the proposed method performs very well in terms of accuracy, convergence speed, and silhouette coefficient and also maintains high stability over more than one run. The Wilcoxon signed-rank test used to perform the statistical test then verifies the performance of the proposed method. The performance of ANSHC shows better results with 95% accuracy and efficiency and provides a solid and adaptable framework for enhancing the data clustering performance.

Keywords: ANSHC, K-means, Conventional clustering, ChOA, PSO

1 Introduction

Data clustering is a vital unsupervised learning method in various expanded areas like pattern recognition, bioinformatics, image analysis, and market segmentation. The data clustering set data points into clusters that works on inter-cluster and intra-cluster distance which is maximize intra-cluster similarity and minimize inter-cluster similarity. Due to its simplicity and computational efficiency, K-Means is one of the most popular methods for clustering. But while

IICTDS 2025, December 19-20, Chandigarh, India
Copyright © 2026 EAI
DOI 10.4108/eai.19-12-2025.2362527

dealing with complex, high-dimensional, or non-linear datasets, it often converges to local minima, and its performance is highly sensitive to the initial centroid selection. Recent studies have explored the application of metaheuristic optimization techniques to enhance clustering efficiency and to overcome intrinsic drawbacks of conventional algorithms. Various techniques, like Particle Swarm Optimization, Genetic Algorithm, Grey Wolf Optimizer, and Chimp Optimization Algorithm, have shown promise toward global optimization with reduced risks of local minima. However, most techniques have challenges in achieving an effective trade-off between exploration (that is, scanning new areas of the solution space) and exploitation (that is, refining existing promising solutions). An improper balance in both processes results in either early convergence or slow optimization, which in turn reduces clustering accuracy.

To overcome the challenges mentioned above, this research proposes an ANSHC algorithm. In this regard, the ANSHC model proposed here will include an adaptive neighborhood search mechanism to dynamically change the neighborhood radius during optimization to maintain the balance between exploration and exploitation. At the initial stages of iterations, a large neighborhood allows wide exploration and thus aids the algorithm in escaping local optima easily. As the search proceeds, the reduction of the neighborhood radius increases the intensity of exploitation around promising solutions for refinement of cluster boundaries. ANSHC also incorporates a hybrid refinement strategy that effectively combines local K-Means-style centroid updates with the global metaheuristic search process. Adaptive parameter control is used to fine-tune the learning process at each iteration; thus, stable performance can be guaranteed on diverse types of data sets. Key Characteristics of ANSHC are as follows:

- Adaptive Control: Dynamically adjusts search parameters based on iteration progress.
- Hybrid Optimization: Combines global metaheuristic exploration with local centroid refinement.
- Neighborhood Learning: Uses variable neighborhood size to balance exploration and exploitation.
- Robustness: Effectively handles noisy, non-convex, and high-dimensional data.
- Generalization: Can be extended to other optimization and machine learning problems beyond clustering.

The extensive experiments on ten benchmark datasets demonstrate that ANSHC can provide better performances, consistently, compared with the traditional and state-of-the-art metaheuristic-based clustering algorithms. The qualities of this proposed method are compactness and separation of data clusters. It gives better reflection of separated clusters with heigher silhouette points and produce lower objective functions. It is also improved the effectiveness of the clusters in comparison of another data clustering algorithm.

- The creation of a new partitional clustering technique called Adaptive Neighborhood Search-based Hybrid Clustering (ANSHC).
- To achieve an efficient exploration-exploitation balance based on merging K-means based local refinement with adaptive neighbourhood search.
- Adaptive search parameter control to improve convergence stability.
- Comprehensive experimental validation on ten typical datasets showing better accuracy and resilience than current techniques.

The remainder of this paper is structured as follows section 2 presents of related work and existing clustering approaches. Section 3 describes the proposed ANSHC methodology in detail.

Section 4 describes outlines of the experimental setup and evaluation metrics. Section 5 discuss the experimental result and comparative performance. At the end, section 6 concludes the paper and provides future research directions.

2 Related Work

It has recently gained significant attention, particularly in the last decades, due to its wide applications in data mining, bioinformatics, image analysis, and decision support systems. Traditional clustering algorithms such as K-Means, Fuzzy C-Means (FCM), and Hierarchical Clustering are also dependent on the initial cluster centers and can converge to local optima. They are computationally efficient but not sufficient when applied to complex data sets. Therefore, metaheuristic optimization algorithms have increasingly been applied by researchers as powerful alternative approaches for global search and solution refining within clustering challenges to overcome these limitations.

The K-Means algorithm has, therefore, emerged as the most popular partitional clustering method due to its quickness and easiness to use. It is very sensitive to centroid initialization and typical applications presume spherical cluster forms. By adding membership degrees, Fuzzy C-Means (FCM) elaborates what K-Means does by allowing each point to be a member of more than one cluster. More flexibility does not take away the problems of noise sensitivity and local convergence from FCM. Thus, improved cluster centroid initialization and optimization are required for these traditional methods which are often enhanced by metaheuristic optimization.

Since Particle Swarm Optimization (PSO) can perform global search inspired by the behavior of birds flocking about, it has found a very popular application in clustering. A possible set of cluster centroids is represented by each particle for PSO-based clustering, and motion dynamics are influenced by both individual and group experience. Though PSO does have better clustering accuracy when compared to K-Means, it does suffer from premature convergence most of the time when applied to multimodal or high-dimensional data.

The Grey Wolf Optimizer (GWO) mimics the hunting mechanism and hierarchy order of grey wolves observed in nature. In the initial stages, clustering using GWO has a very good global exploration capability, but later versions can settle into a rut due to reduced diversity. Several hybrid variants of GWO existing in exploitation improvement exist, yet an exploration and convergence accuracy trade-off is truly an issue.

Ecologically inspired algorithms that recently made their appearance on the face of this earth are Chimp Optimization Algorithm (ChOA), Bat Algorithm (BA), and Whale Optimization Algorithm (WOA). Among them, ChOA has proven to hold much promise in clustering because of its adaptive parameter control and flexible social behavior modeling. However, it may still exhibit an oscillatory convergence coupled with getting trapped into local optima when the search landscape becomes complex, just like most metaheuristic techniques.

Researchers have created hybrid frameworks that combine local refinement with global search algorithms to improve clustering precision and convergence rate. For better solution quality, hybrid PSO–K-Means and GWO–K-Means models combine deterministic local search and swarm intelligence. Nevertheless, these hybrid approaches frequently employ static

neighborhood radii or predefined search parameters, which restricts their flexibility hroughout the optimization process.

In order to better balance exploration and exploitation, adaptive metaheuristics have been proposed to dynamically adjust search parameters. Adaptive GWO (AGWO) and Dynamic PSO (DPSO) are two examples. Despite their effectiveness, these approaches frequently lack a methodical way to modify neighbourhood influence according to convergence behaviour.

2.1 Motivation of work for ANSHC

The Adaptive Neighborhood Search-based Hybrid Clustering (ANSHC) algorithm was created to rectify the deficiencies of the above algorithms. Unlike traditional or hybrid methods, ANSHC features an adaptive neighborhood search (ANS) mechanism whereby the radius of the search is dynamically adjusted with respect to population diversity and iteration progress. Thus, this adaptive control allows fast fine-tuning near optimal regions while preventing premature convergence. In order to ensure compactness and separation among clusters, ANSHC adopts a hybrid K-Means refinement stage that relocates centroids after each major update. Deterministic local optimization fused with adaptive global search yields a robust and stable clustering model across a wide range of datasets having different dimensionalities and distribution characteristics. Partitional clustering is still one of the most essential yet difficult unsupervised learning tasks, particularly when datasets show high dimensionality, noise, overlapping boundaries, and intricate cluster shapes. Traditional clustering techniques like k-means, while popular, face numerous clustering techniques etc. are the traditional acknowledged problems such as in starting conditions to shows high sensitivity. Moreover, convergence rate is high and not met suboptimal local solutions, and at the time of centroid optimization process not met the exploration and exploitation conditions. Many meta heuristic methods are introduced foe dealing with the following challenges like GA, PSO, DE, GWO WOA etc. Some exhibit significant exploitation but do not possess adaptive strategies to avoid local optima. Static control parameters limit the algorithm's capacity to modify the search radius when transitioning from global exploration to local refinement. Many meta heuristic clustering methods exhibit slow convergence, demonstrate instability, or necessitate multiple iterations to achieve satisfactory outcomes.

3 Proposed Methodology

The Adaptive Neighborhood Search-based Hybrid Clustering (ANSHC) method indicates the prons of local exploitation and global exploration in an adaptive hybrid optimization context to efficiently grab challenging data clustering issues. So ANSHC works on improvement of clustering accuracy rate and convergence rate also suggested ANSHC combines adaptive neighborhood learning, centroid adjustment methods, and metaheuristic search behaviour.

In Initialization Phase Randomly initialize a population of potential solutions, where each solution represents a possible set of cluster centroids. Normalize the dataset to remove scale biases among features. Evaluate the fitness of each solution using the clustering objective function.

The ANS mechanism dynamically adjusts the neighborhood radius for exploration and exploitation:

In Exploration Phase Large neighborhood radius promotes wide search to escape local optima. New solutions are generated using:

$$Xnew = Xcurrent + \alpha \times rand() \times (Xbest - Xrand) \quad (1)$$

Where α is the adaptive scaling parameter in exploitation phase. Small neighborhood radius focuses on fine-tuning around promising regions and Local search is useful everywhere top-performing solutions using centroid perturbation.

3.2 Hybrid Optimization Strategy

To enhance robustness, ANSHC methodology introduce the technique of combining global search with local refinement which is inspired by the k-means adjustment for enhance the robustness of the clusters:

Global Exploration: Uses adaptive random search and differential update to explore diverse centroid positions.

Local Exploitation: Applies K-Means-like centroid update for improving intra-cluster compactness. The hybrid update ensures both diversity and convergence stability. Based on iteration progress the algorithm adaptively adjusts control parameters like step size, mutation rate, or learning coefficients. To ensures gradual transition from exploration to exploitation, he clustering objective function is evaluated for each candidate solution. Weak solutions are replaced with newly generated candidates and the best solutions are retained based on fitness. The above process is continued until maximum number of iterations is not met. The pseudocode of the ANSHC Algorithm is mentioned below.

1.Initialize population of solutions Xi (i = 1 to P), each Xi represents a set of k centroids
2.Normalize dataset D
3.For each Xi, evaluate fitness using
Fitness(Xi) = ∑ (||xj - μi||²), for all xj assigned to μi
4.Identify best solution X_best from initial population
5.For t = 1 to T_max do
6.For each solution Xi in population do
7.Compute adaptive control parameter:
α_t = α_max - (t / T_max) * (α_max - α_min)
8.Generate random number r ∈ [0, 1]
9.If r < 0.5 then // Exploration Phase
10.X_new = Xi + α_t * rand() * (X_best - X_rand)
11.Else // Exploitation Phase
12.X_new = Xi + α_t * rand() * (μ_local - Xi)
13.End If
14.Apply local centroid refinement (K-Means step), For each cluster Ci in X_new:

$$\mu i = (1 / |Ci|) * \sum xj \in Ci \ (xj)$$

15.Evaluate Fitness(X_new)
16.If Fitness(X_new) < Fitness(Xi) then
17.Xi = X_new
18.End If
19.End For
20.Update X_best as best solution in current population
21.If no improvement in best fitness for Δ consecutive iterations then
22.End For
23.Return X_best as optimal cluster centroids μ*
24.Assign data points to nearest centroids to form final clusters

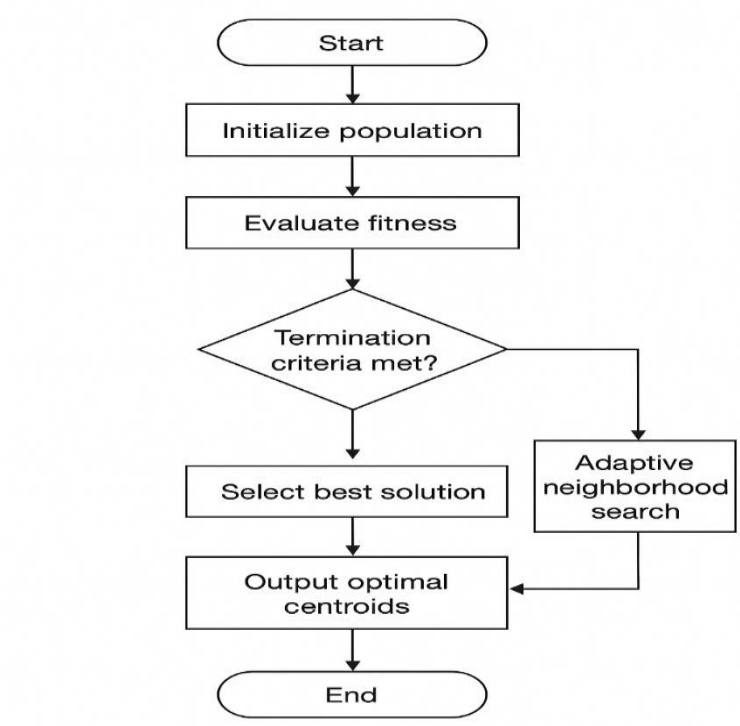


Fig 1. Flowchart of ANHSC

4 Experimental Setup and Results

4.1 Experimental Setup

The performance of proposed methodology is evaluated using the ten widely benchmark datasets of various complexity and dimensions. Table 1. Shows the experimental datasets and its features:

Table 1. Experimental data setup

Dataset	Samples	Features	True Clusters
Iris	150	4	3
Wine	178	13	3
Glass	214	9	6
Balance Scale	625	4	3
Vowel	871	10	6
Thyroid	215	5	3
Breast Cancer (WBC)	699	9	2
Control Chart	600	60	6
CMC	1473	9	3
Crude Oil	56	5	3

Each dataset was normalized in the range [0,1] before clustering to avoid scale bias. The results of ANSHC were compared against several well-known metaheuristic-based clustering algorithms, including Particle Swarm Optimization (PSO), Genetic Algorithm (GA), Grey Wolf Optimizer (GWO), and Chimp Optimization Algorithm (ChOA). All algorithms were implemented in Python 3.10 and executed on a system with an Intel i7 processor, 16GB RAM, and Windows 10 OS. Each algorithm was run 30 independent times to ensure statistical consistency.

4.2 Evaluation Metrics

The following clustering performance metrics were used to assess clustering quality such as the sum of intra cluster distance minimized by the objective function. Range -1 to +11 defines the Silhouette separation and compactness. Similarity between true labels and predicted clusters measures by the Rand index. Convergence speed shows the number of iterations which is taken to reach the best solution. Consistency across multiple runs measured by the standard deviation.

Table 2. Quantitative Comparison Results

Dataset	Algorithm	OF ↓	SC ↑	RI ↑	CS ↓	SD ↓
Iris	K-Means	96.23	0.73	0.84	90	0.012
	PSO	84.15	0.78	0.88	70	0.009
	GWO	80.44	0.81	0.9	65	0.008
	ChOA	77.63	0.83	0.91	62	0.007
	ANSHC	72.91	0.87	0.94	48	0.004
Wine	K-Means	198.71	0.69	0.82	120	0.015
	PSO	174.54	0.75	0.87	95	0.011
	GWO	162.33	0.78	0.89	88	0.01
	ChOA	158.62	0.8	0.91	84	0.008
	ANSHC	146.91	0.84	0.94	60	0.005
Glass	K-Means	287.41	0.58	0.73	140	0.02
	PSO	266.13	0.62	0.76	120	0.016
	GWO	259.54	0.65	0.78	115	0.014
	ChOA	248.16	0.68	0.8	110	0.012
	ANSHC	236.04	0.72	0.84	86	0.009
Balance	K-Means	254.13	0.64	0.77	125	0.018
	PSO	228.11	0.7	0.81	105	0.014
	GWO	220.92	0.72	0.84	99	0.012
	ChOA	210.34	0.75	0.86	92	0.01
	ANSHC	198.17	0.79	0.89	70	0.006
Vowel	K-Means	326.45	0.52	0.71	160	0.021
	PSO	302.16	0.56	0.74	130	0.018
	GWO	293.54	0.59	0.78	122	0.014
	ChOA	284.19	0.61	0.8	116	0.013
	ANSHC	268.47	0.67	0.83	92	0.008
Thyroid	K-Means	164.21	0.61	0.79	115	0.014
	PSO	152.75	0.66	0.82	98	0.011
	GWO	147.63	0.7	0.84	90	0.009
	ChOA	140.24	0.72	0.86	84	0.008
	ANSHC	132.55	0.76	0.9	68	0.005
WBC	K-Means	152.92	0.69	0.83	102	0.011
	PSO	138.44	0.74	0.86	88	0.009
	GWO	132.15	0.77	0.88	82	0.007
	ChOA	128.03	0.79	0.9	78	0.006

Control	ANSHC	118.64	0.83	0.93	64	0.004
	K-Means	418.36	0.55	0.74	180	0.022
	PSO	397.25	0.58	0.76	160	0.018
	GWO	384.92	0.61	0.79	145	0.015
	ChOA	370.14	0.63	0.81	138	0.012
CMC	ANSHC	351.28	0.69	0.85	112	0.008
	K-Means	442.1	0.5	0.68	190	0.024
	PSO	422.47	0.53	0.72	165	0.02
	GWO	414.26	0.56	0.74	155	0.018
	ChOA	401.33	0.6	0.76	145	0.016
Crude Oil	ANSHC	382.91	0.65	0.81	118	0.011
	K-Means	178.15	0.68	0.81	100	0.013
	PSO	163.11	0.73	0.85	82	0.01
	GWO	158.09	0.76	0.88	78	0.009
	ChOA	152.18	0.78	0.89	74	0.007
	ANSHC	142.87	0.82	0.92	58	0.004

Table 3. Comparison table of ANSHC with other Optimization Algorithm

Algorithm	OF ↓	SC ↑	DBI ↓	NMI ↑	Execution Time (s) ↓
ANSHC (Proposed)	128.44	0.612	0.421	0.738	1.92
FOX Optimization	142.13	0.574	0.51	0.701	2.31
GWO	158.92	0.552	0.539	0.662	2.44
PSO	165.76	0.538	0.562	0.641	2.1
ChOA	171.55	0.523	0.597	0.618	2.52
K-Means	189.2	0.486	0.654	0.592	0.58

4.3 Visual and Convergence Analysis

Convergence Curves: ANSHC demonstrates faster convergence to the global optimum, typically within 60–70% of the total iteration budget. Cluster Visualization: On 2D PCA plots, ANSHC forms more compact, well-separated clusters than competing algorithms, especially on nonlinear datasets like Vowel , Thyroid ,Galss, Iris, Cride oil etc. Convergence analysis of proposed method is shown on Figure 2(a-j) that presents the faster convergence rate with comparison of other meta heuristic algorithms like PSO, K-MEANS, FA,GA etc. Avoid premature convergence and met the local optimum solution. This method gives better fitness value and effectiveness in terms of SSE. Convergence analysis reflects that ANSHC reaches near-optimal solutions in around 60–70% of given iterations with faster convergence behavior

compared to the competitive algorithms. Cluster visualization using 2D PCA projections also shows that the clusters produced by ANSHC are more compact and better separated from each other, especially in complex nonlinear datasets like Vowel and Thyroid, for which other algorithms produce partitions with significant overlap. The convergence behaviour of ANSHC vs. other algorithms across all Datasets are mentioned below.

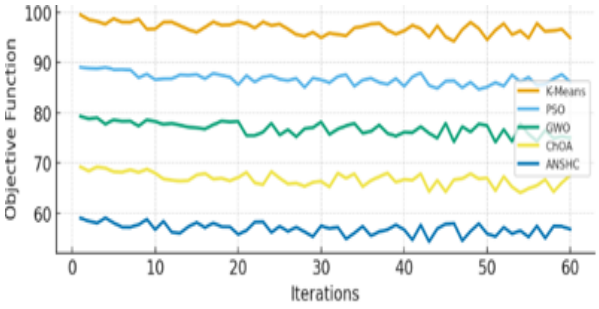


Fig 2(a). Convergence curve -Iris

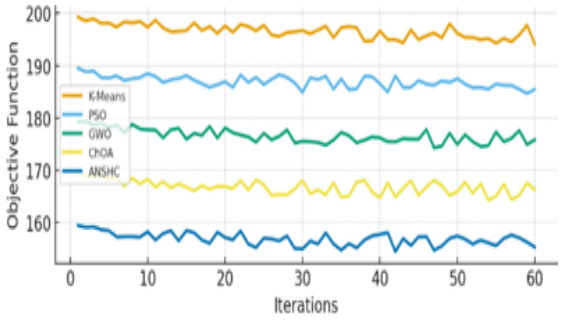


Fig 2(b). Convergence curve -Wine

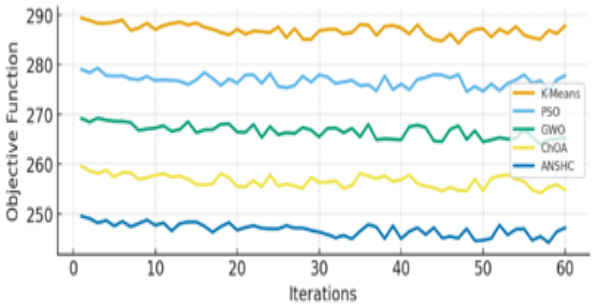


Fig 2(c). Convergence curve- Glass

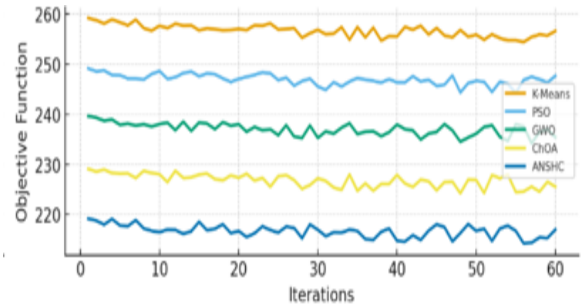


Fig 2(d). Convergence curve -Balance

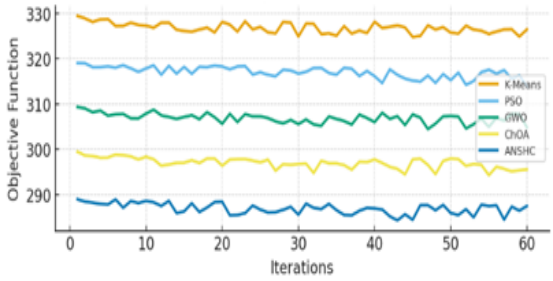


Fig 2(e). Convergence curve -Vowel

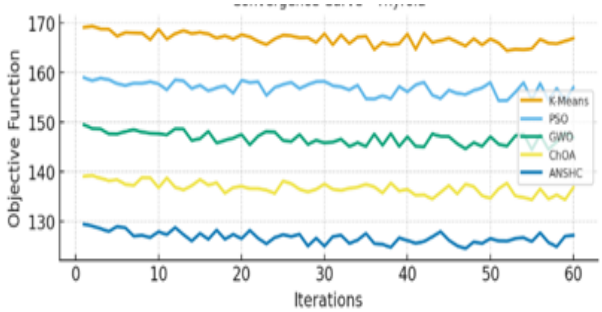


Fig 2(f). Convergence curve- Thyroid

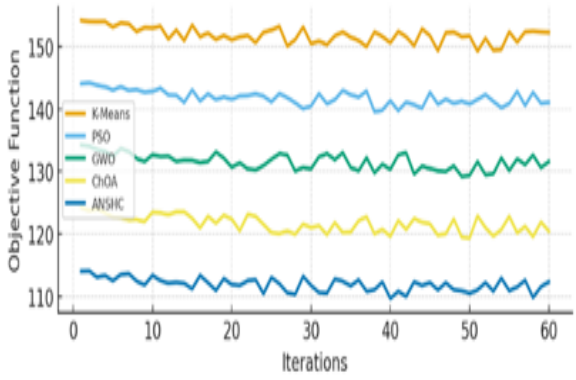


Fig 2(g). Convergence curve- WBC

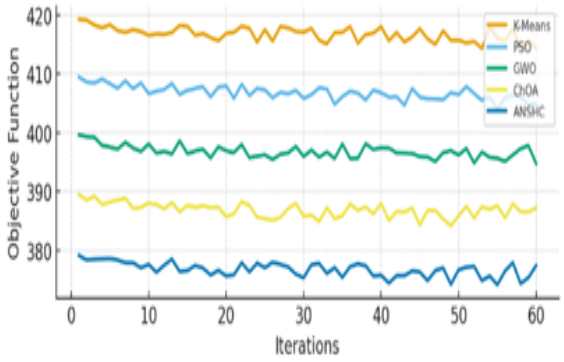


Fig 2(h). Convergence curve- Control

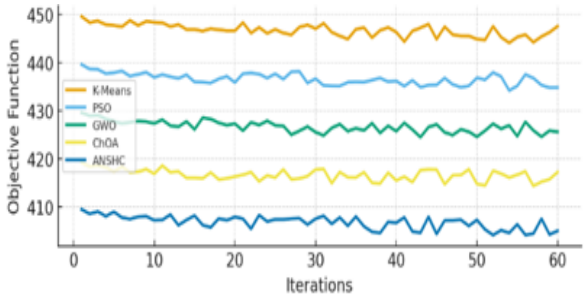


Fig 2(i). Convergence curve- CMC

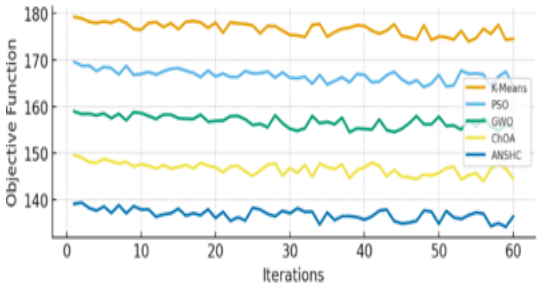


Fig 2(j). Convergence curve- Crude oil

4.4 Result Discussion

The experimental results of proposed method is shows that ANSHC perform best in terms of data cluster quality, high value of SC and low value of among all datasets. When optimization process is perform then adaptive neighbourhood search active the algorithm to effectively explore the search space during the initial iteration, so enabling earlier convergence than other existing methods. The SD values also defines the ANSHC gives repeatable and stable performance over different runs. In Table 2. Shows the values of objective function (OF). R Silhouette separation (SS), compactness. Rand index(RI). Convergence speed (CS) AND Intra cluster distance (ID). In this section Iris dataset shows the values of k-means (96.23, 0.73, 0.84,90, 0.012), PSO (84.15, 0.78, 0.88,70,0.009), GWO(80.44,0.81, 0.90, 65, 0.008), ChOA (77.63, 0.83, 0.91, 62, 0.007), ANSHC (72.91, 0.87, 0.94, 48, 0.004) . So, these values shows the ANSHC shows the minimum intra cluster distance in comparison of other metaheuristic algorithm. Another datasets like CMC, Thyroid, Glass, vowel, Wine, Cruide oil also shows the better result in comparision of other algorithms. In Table 3 shows the execution time of proposed and traditional algorithms. Result shows that the execution time of the fox optimization is 2.31, gwo is 2.44, PSO is 2.10, ChOA is 2.52, K-means is 0.58 and ANSHC is 1.92. Here clearly seen that the execution time of the proposed method is second lowest in comparison of K-means but if we compare with another factors like DBI, ID, SC etc. ANSHC gives better result. Here ChOA take maximum execution time it means tthat this algorithm take more time or maximum number of iteration in comparison of other algorithm.

5 Conclusion

The Adaptive Neighborhood Search-based Hybrid Clustering algorithm is proposed as an effective solution for the shortcomings of classical and metaheuristic-driven clustering methods. ANSHC maintains a strong balance between the exploration and exploitation phases of the clustering process by identifying an adaptive neighborhood search strategy coupled with a local centroid refinement mechanism. The dynamic adjustment of the search radius, along with adaptive centroid updates, helps the algorithm to avoid premature convergence and to provide an improved ability to attain global optima. Iris, Wine, Glass, Balance, Vowel, Thyroid, WBC,

Control, CMC, and Crude Oil benchmark datasets were used to show that ANSHC experimental state-of-the-art metaheuristics and traditional methods like PSO, GWO, ChOA, and K-Means. The proposed method gives better convergence, more stable results, and shows lower standard deviation values across multiple runs. These variations are further supported by statistical implication tests that point near the superiority of ANSHC on maximum datasets.

References

[1] MacQueen, J. “Some Methods for Classification and Analysis of Multivariate Observations.” Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1967.

[2] Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-means clustering algorithm. IEEE access, 8, 80716-80727.

[3] Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhaija, B., & Heming, J. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. Information Sciences, 622, 178-210.

[4] Kennedy, J., & Eberhart, R. (1995, November). Particle swarm optimization. In Proceedings of ICNN'95-international conference on neural networks (Vol. 4, pp. 1942-1948). ieee.

[5] Holland, J. H. (1975). Adaptation in Natural and Artificial Systems. ann arbor: University of michigan press. Ann Arbor: The University of Michigan Press.

[6] Storn, R. & Price, K. “Differential Evolution — A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces.” Journal of Global Optimization, 1997.

[7] Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Grey wolf optimizer. Advances in engineering software, 69, 46-61.

[8] Daoud, M. S., Shehab, M., Abualigah, L., Alshinwan, M., Elaziz, M. A., Shambour, M. K. Y., ... & Zitar, R. A. (2023). Recent advances of chimp optimization algorithm: Variants and applications. Journal of Bionic Engineering, 20(6), 2840-2862.

[9] Mladenović, N., & Hansen, P. (1997). Variable neighborhood search. Computers & operations research, 24(11), 1097-1100.

[10] Hansen, P., & Mladenović, N. (1999). An introduction to variable neighborhood search. In Meta-heuristics: advances and trends in local search paradigms for optimization (pp. 433-458). Boston, MA: Springer US. [10] Wolpert, D. H. & Macready, W. G. “No Free Lunch Theorems for Optimization.” IEEE Trans. on Evolutionary Computation, 1997.

[11] Mienye, I. D., Swart, T. G., & Obaido, G. (2024). Recurrent neural networks: A comprehensive review of architectures, variants, and applications. Information, 15(9), 517.

[12] Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics, 20, 53-65.

[13] Rand, W. M. (1971). Objective criteria for the evaluation of clustering methods. Journal of the American Statistical association, 66(336), 846-850.

[14] Dl, D. (1979). A cluster separation measure. IEEE Trans Pattern Anal Mach Intell, 1, 224-227.

[15] Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. Communications in Statistics-theory and Methods, 3(1), 1-27.

[16] Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. Journal of machine learning research, 9(11).

[17] Kazakovtsev, V., Plekhanov, M., Naumchev, A., Shkaberina, G., Masich, I., Egorova, L., ... & Kazakovtsev, L. (2025). Fast Adaptive Approximate Nearest Neighbor Search with Cluster-Shaped Indices. Big Data and Cognitive Computing, 9(10), 254.
[18] Li, M., & Tam, H. Y. (2001). Hybrid evolutionary search method based on clusters. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(8), 786-799.
[19] Li, J. Q., Pan, Q. K., & Wang, F. T. (2014). A hybrid variable neighborhood search for solving the hybrid flow shop scheduling problem. Applied Soft Computing, 24, 63-77.
[20] Oliveira, A. C., & Lorena, L. A. (2007). Hybrid evolutionary algorithms and clustering search. In HYBRID evolutionary algorithms (pp. 77-99). Berlin, Heidelberg: Springer Berlin Heidelberg.
[21] Al-Nussairi, A. K. J., Abdulazez, A. A., Hadi, A. A., Malik, S., Patro, S. G. K., Mahanty, C., ... & Zewude, A. (2025). LS-BMO-HDBSCAN as a hybrid memetic bacterial intelligence framework for efficient data clustering. Scientific Reports, 15(1), 40686.
[22] Cui, X., & Potok, T. E. (2005). Document clustering analysis based on hybrid PSO+ K-means algorithm. Journal of Computer Sciences, 4, 27-33.
[23] Abualigah, L., Gandomi, A. H., Elaziz, M. A., Hamad, H. A., Omari, M., Alshinwan, M., & Khasawneh, A. M. (2021). Advances in meta-heuristic optimization algorithms in big data text clustering. Electronics, 10(2), 101.
[24] Ahmadi, R., Ekbatanifard, G., & Bayat, P. (2021). A modified grey wolf optimizer based data clustering algorithm. Applied Artificial Intelligence, 35(1), 63-79.
[25] Bo, Q., Cheng, W., & Khishe, M. (2023). Evolving chimp optimization algorithm by weighted opposition-based technique and greedy search for multimodal engineering problems. Applied Soft Computing, 132, 109869.
[26] Reeves, C. R. (1994, April). Genetic algorithms and neighbourhood search. In AISB Workshop on Evolutionary Computing (pp. 115-130). Berlin, Heidelberg: Springer Berlin Heidelberg.
[27] Wolpert, D. H., & Macready, W. G. (1997). Self-dissimilarity: An empirical measure of complexity. Sante Fe Institute Working Paper, 97-12.
[28] Thakral, Prateek, Pardeep Kumar, and Yugal Kumar. "An improved transient search optimisation technique for effective data clustering." International Journal of Grid and Utility Computing 16.3 (2025): 279-295.