Whale Optimization Algorithm for Data Clustering

David
STMIK Pontianak, Pontianak, Indonesia
DavidLiauw@gmail.com

Muhammad Qadafi Khairuzzaman
STMIK Pontianak, Pontianak, Indonesia
m.qadafi.k@gmail.com

Gusti Syarifudin
STMIK Pontianak, Pontianak, Indonesia
gus_wet@yahoo.com

Abstract— Issues relating to clustering today are computational techniques, optimization and performance of clustering algorithms. In this research, a metaheuristic grouping method of Whale Optimization Algorithm with modification to weight changes based on food hunting method and behaviour of humpback whales was proposed. A total of ten datasets obtained from the learning repository of UCI machine were used to evaluate performance of the proposed algorithm compared to other clustering algorithms. Whale Optimization Algorithm can provide optimal solutions and more stable clustering results because there is no dependence on initial cluster center initialization. Moreover, clustering using Whale Optimization Algorithm produces clusters which are better than clusters using Ant Colony Optimization, Isodata, and Forgy. This can be seen from silhouette coefficients, time examination, variance examination and sum of square error of a number of test datasets obtained. Overall, results show that clusters of Whale Optimization Algorithm and the other three methods have nearly identical variance, and each cluster produces high intra-class similarity and low inter-class similarity.

Keywords—Clustering, Whale Optimization Algorithm, Silhouette Coefficients, Datasets, Objective Functions

I. INTRODUCTION

Data have essential roles in all aspects of human life. It is understandable that analyzing data, extracting useful information from data, and transforming data are primary tasks of data mining. Clustering is another task of this process. It is a form of unsupervised learning of similar objects placed in the same cluster [1].

Clustering refers to a data processing technique used to find hidden patterns in the dataset [1]. Clustering has been widely applied in various fields including data mining, pattern recognition, decision making, machine learning, and image segmentation [2]. The process of finding data patterns is conducted by grouping data into clusters. In other words, those with similarity are placed into a certain cluster, while those with dissimilarity are located in other clusters [3]. A way taken to find out similarity levels of data is through calculation of the distance between data. The smaller the distance of data indicates a higher similarity level. On the other hand, the greater the distance of data indicates a lower similarity level [3].

Clustering is a powerful technique in data mining involving identification of homogeneous groups of objects based on attribute values. Metaheuristic algorithms (optimization of particle flocks, artificial bee colonies, genetic algorithms, and differential evolution) are currently powerful grouping methods [4].

Metaheuristic is an algorithm that can solve complex optimization problems provided that exact algorithms are encountered. Heuristic method is, however, a method used to find the most feasible solutions to a problem [4]. When searching efficient, comprehensive solutions, metaheuristic methods apply mechanisms imitating social behavior or strategies existing in nature [4]. A metaheuristic algorithm has a searching speed for optimal solutions which are better than traditional methods [5]. Compared to a heuristic method, it is better since it always obviates local optima solutions. Despite no guarantee, a well-built metaheuristic method can provide a nearly optimal solution [4] [5].

Whale Optimization Algorithm (WOA) is one of the metaheuristic methods [6]. It is an evolutionary type of algorithm inspired by humpback whales [6]. At first, WOA was made to solve optimization problems. Nonetheless, in the recent few years, it has been applied to solve various clustering-related problems [7][8]. Becoming the development of Particle Swarm Optimization, this algorithm can provide more stable clustering results due to no dependence on cluster center initialization [8]. Despite having a competitive convergence level, WOA is weak in its convergence speed tending to be slow when approaching the optimum solutions [6].

In this study, a recent metaheuristic grouping method of WOA with modification to weight changes based on food hunting method and behaviour of humpback whales was proposed. After detailed formulation and elaboration of implementation, the proposed algorithm and other well-known clustering algorithms were compared. It was tested by using ten datasets obtained from the learning repository of UCI machine.

II. PROPOSED METHOD

The method proposed to solve the clustering problem was a metaheuristic and population-based algorithm. It was noted that several whale agents performing as a solution were created at random. This candidate solution found a cluster center. Cluster solutions were found when the search whale agent had the best objective function value.

The main contribution of the proposed method was to use WOA to optimize objective functions so that grouping performance was achievable. To reach the goal, such functions representing whale agents in WOA with dimensions determined by the number of clusters were introduced. The whales moved through the search space using algorithm rules. Addtionally, this process continued until the convergence criteria were met.

This research began with literature review concerning WOA and its use classified as swarm intelligence. The intention was to solve clustering problems. The following step was collecting data by preparing datasets needed in algorithm. Next, a method was designed by WOA designer to cluster data. Another stage was on implementation of the clustering method. Here, the proposed design was applied on program rows. Finally, the program was used to cluster datasets obtained from the UCI Machine Learning Dataset.

In the examination phase of each dataset, each proposed clustering of WOA was compared to the quality of the clustering scheme with three other algorithms, namely Forgy, Isodata, and Ant Colony Optimization Clustering (ACOC) using testing criteria. The clustering examination was performed for each dataset and each different clustering method. Values viewed from duration of clusters, Sum of Square Error (SSE), and silhouette coefficients were additionally taken.

A. Whale Optimization Algorithm (WOA)

WOA refers to one of the optimization algorithms developed by Sayedali Mirjalili and Andew Lewis [6]. A whale is one of the largest mammals in the world. There are many types of whales. One of them is a humpack whale [9]. This animal is very intelligent as they have spindle cells in its brain. Humpback whales are also known from their pectoral fins reaching 4.6 m in length. Their scientific name is, *Megaptera Novaeangliae*. Their long fins make them have high maneuverability and ability to slow down or even go backwards [9]. Whales live in groups and can communicate with each other. One of the interesting things about humpback whales is their hunting method. Their favorite food is krill or groups of small fish.

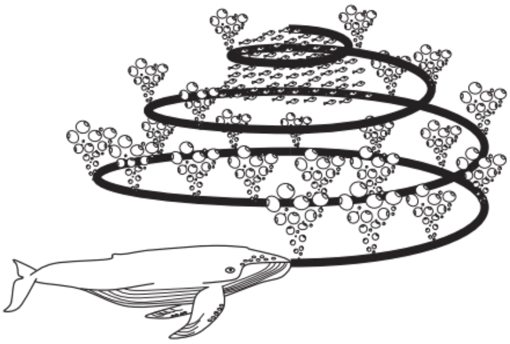

Fig. 1. Bubble-net technique of humpback whales [6]

Some hunting methods include the use of air bubbles in herding, collecting or disorienting the fish. A very complex variant called "mosquito net" is uniqueness of humpback whales [10]. This technique is often applied in groups with defined roles to make disturbance and scare before they attack their prey when it is near the surface. A group of whales will surround it while removing bubbles [10]. The diameter of circumference becomes smaller and narrows the target space. Bubbles released become nets used to trap the prey. One of the whales swims up through bubble nets and catches it after all [10].

Three mathematical models used in the technique of bubble-net feeding of this algorithm are as follows:

a. Surrounding the Prey

At this stage, humpback whales recognize the location of the target prey. Next, they surround it. An optimal position of the search space is unknown a priori. Thus, WOA assumes that the best solution is to target prey surrounded. In other words, the position at this time is nearly optimal. After the best search whale is defined, the other whales attempt to change their positions and approach it. This behavior can be illustrated with the following equations:

$$\vec{D} = |\vec{C}\vec{X}(t) - \vec{X}(t)| \qquad (1)$$

$$\vec{X}(t+1) = \overrightarrow{X^*} - \vec{A}\vec{D} \qquad (2)$$

To give elaboration, t shows the current iteration, $\vec{A}$ and $\vec{C}$ are the vector coefficients, $\overrightarrow{X^*}$ is the position vector of the best solution currently obtained, X is the position vector to be updated, and | | is an absolute value. The above equations indicate that X* must be updated in each iteration before finding a better solution. Vectors $\vec{A}$ and $\vec{C}$ are computed as follows:

$$\vec{A} = 2\vec{a}\vec{r} - \vec{a} \qquad (3)$$

$$\vec{C} = 2\vec{r} \qquad (4)$$

$\vec{a}$ is the linear reduction from 2 to 0 during the iteration in both the exploration and exploitation (to be explained further in the following part) and $\vec{r}$ is a random vector in [0,1].

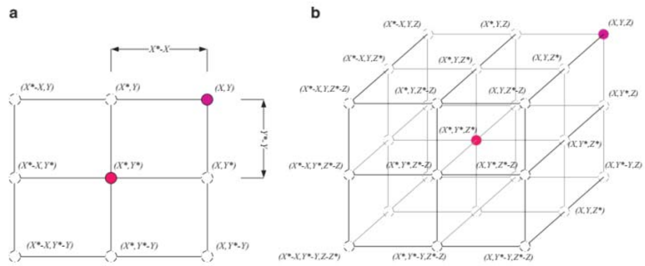

Fig. 2. Vector positions of whale movement on 2D and 3D problems, and possibilities of whale movement to the nexxt position (X* is the currently best position) [6].

Figure 2 (a) illustrates the movement of whales from Equation (2) on a two-dimensional (2D) problem. Position (X,Y) of a search whale can be updated based on the current best position (X*,Y*). The different positions around the best whale can be achieved by considering the position of the other whales at this time and make adjustment to vector values $\vec{A}$ and $\vec{C}$. The possible movement of the position of the search whale on a three-dimensional (3D) problem is illustrated in Figure 2 (b). It should be noted that by defining random vector ($\vec{r}$), the search whale can reach any position in the scope of the search space as shown in Figure 2. Thus, Equation (2) allows several other search whales to change their positions to approach the currently best whale position so that the movement of the group of whales can surround the prey.

A similar concept can be used as search space with n dimension [12]. Then, a group of search whales move to hyper-cubes surrounding the best solution that has been obtained in the iteration to t.

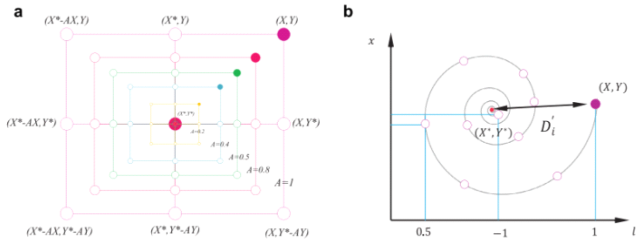

Fig 3. Bubble-net mechanism where X* is the best position at iteration of t. a) shrinking and encircling mechanism b) spiral updating position [6].

b. Bubble-net Attack / Exploitation Phase

Mathematically, this exploitation phase can be represented in the following two stages:

Shrinking and encircling mechanism

This mechanism is obtained by reducing the value from $\vec{a}$ in Equation (3). It should be noted that changes in range value $\vec{A}$ decreases as $\vec{a}$ decreases. In other words, $\vec{A}$ is the random value at interval [-a, a], where a decreases linearly from 2 to 0 in each iteration. By setting random value $\vec{A}$ at [-1,1], the latest position of the search whale can change according to the interval of the range to the position of the best solution. In Figure 3 (a), it is shown that in a 2D problem, the position movement (X,Y) against (X*,Y*) can be achieved in the range $0 \leq A \leq 1$.

Spiral updating position

As shown in Figure 3 (b), the approach is through calculation of the position of humpback whales (X, Y) to the position where there is prey, namely (X*, Y*). The spiral equation made between the position of whales and the prey imitating the spiral humpback whale movement is as follows:

$$\vec{X}(t+1) = \vec{D'}.e^{bl}.\cos(2\pi l) + \vec{X^*}(t) \qquad (5)$$

Where:

$\vec{D} = |\vec{X^*}(t) - \vec{X}(t)|$, showing the distance of whales to the target (the best solution). *b* is a constant used to define a spiral, logarithmic form and l is a random number between [-1,1].

It should be noted that humpback whales swim in curved circles and along spiral paths simultaneously. To model this simultaneous behavior, it is assumed that there is a 50% chance of humpback whales to choose shrinking and encircling mechanism or spiral updating position to update the position during the iteration. The model is as follows:

$$\vec{X}(t+1) = \begin{cases} \vec{X^*}(t) - \vec{A}\vec{D} \;\; if \; p < 0.5 \\ \vec{D}.e^{bl}.\cos(2\pi l) + \vec{X^*}(t) \;\; if \; p \geq 0.5 \end{cases} \qquad (6)$$

Where p is a random number at interval $0 \leq p \leq 1$ or [0,1].

Search for prey / Exploration

The same approach based on vector variation $\vec{A}$ can be used at the stage of searching for prey or the exploration phase. In fact, humpback whales search the prey at random by looking at each other's position. Hence, random value $\vec{A}$ that is greater than 1 or less than -1 is used to force the search humpback whale to move away from the reference fish. Unlike the exploitation phase, in this exploration phase, the position of the search humpback whale *i* can adjust to the position of the search humpback whale selected at random, not of the best search humpback whale that has been found. This phase is performed if $|\vec{A}|$ >1. Here, WOA performs a global search. The mathematical models are as follows:

$$\vec{D} = |\vec{C}.\overrightarrow{X_{rand}} - \vec{X}| \qquad (7)$$

$$\vec{X}(t+1) = \overrightarrow{X_{rand}} - \vec{A}.\vec{D} \qquad (8)$$

Where $\overrightarrow{X_{rand}}$ is a random vector position selected based on number of humpback whale population.

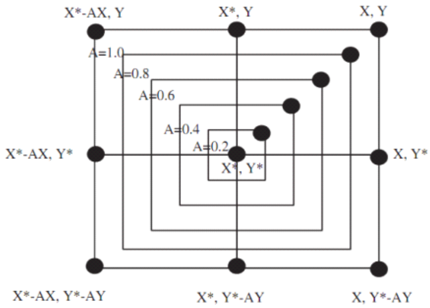

Fig 4. Exploration mechanism of WOA where X* is the best position selected at random [6].

WOA starts with a random, optimum value. In each iteration, the search whale *i* moves closer to the best whale position. For exploration or exploitation, parameter *a* decreases linearly from 2 to 0. Random positions are selected when $|\vec{A}| > 1$, while the best solution is determined when $|\vec{A}| < 1$ to update the position of the search humpback whale. Depending on *p* value, WOA can connect spiral or circular movements. In general, it can be written in pseudocode under [6] [8]:

```
Initialize fish population Xi (i=1,2,…,n)
Calculate the fitness function of each search humpback whale
X* = the best position
While (t < maximal iteration)
  For each search whale
    Update values a, A, C, l, and p
    If (p < 0.5)
      If (|A| < 1)
        The phase of encircling prey is appled with Equation (2)
      Else if (|A| > 1)
        Choose random (Xrand)
        The exploration phase is applied with Equation (8)
      End if
    Else if (p ≥ 0.5)
      The exploitation phase is applied with Equation (6)
    End if
  End for
  Examine is there are whales exceeding search space
  Calculate fitness functions
  Update value X* if it is better
  t = t + 1
End while
Return X*
```

B. *Whale Optimization Algorithm for Clustering*

In the context of WOA, a group of humpback whales refer to a number of potential solutions to problem optimization in which each potential solution is called a search humpback whale. The purpose of WOA is to find the position of the humpback whale hunter producing the best evaluation of objective functions given. In this section, the problem of grouping data will be solved with WOA. Inspired by the context of grouping of humpback whales, it is assumed that search whales represent cluster center *k*. Here, *k*, number of clusters, has been predetermined. Each search whale X_i is written as follows:

$$X_i = (z_{i1}, z_{i2}, \ldots, z_{ik}) \quad (9)$$

where z_{ij} refers to the central vector of cluster-*j* and search whale-*i* in cluster c_{ij}. Therefore, a group of humpback whales represent a number of grouping candidates for vectors from datasets. Calculation of distance among clusters is a fitness function used to measure the distance between the cluster center and vector data from the same cluster according to Equations (10) and (11). Based on Equation (11), each object of data is assigned to the nearest center.

$$d(x,z) = \left(\sum_{i=1}^{n}\sum_{k=1}^{k} w_{ik}\|X_i - z_k\|^2\right)^{\frac{1}{2}} \quad (10)$$

$$w_{ik} = \begin{cases} 1 & if\ \|X_i - z_k\|^2 = min_{1\le j\le n}\|X_i - z_k\|^2 \\ 0 & otherwise \end{cases} \quad (11)$$

The following pseudocode of WOA for clustering is applied [8]:

Take sample data
Initialize each search agent and give value k of the cluster center at random
While (t < iteration)
For each search whale i
For each data vector xp
Calculate euclidean distance from xp to all cluster centers
Specify xp to cluster cij through Equation (12):

$$|x_{ij} - z_{ij}| = \min_{c=1,2,\ldots,k} |x_{ij} - z_{kj}| \quad (12)$$

Calculate fitness value by using Equation (13):

$$fitness = \sum_{j=1}^{k}\sum_{i=1}^{n} w_{ij}|x_{ij} - z_{ij}| \quad (13)$$

Where:

$$w_{ij} = \begin{cases} 1\ if\ |x_i - z_{ij}| = \min_{1\le m\le k} |x_i - z_{im}| \\ 0\ else \end{cases} \quad (14)$$

End for
End for
X* = The best search whale
For each search whale
Update value a, A, C, I and p
If p < 0.5 then
If |A| < 1 then
Update search whales through Equation (2)
Else if |A| >= 1 then
Choose search whales at random
Update current search whales through Equation (8)
End if
Else if p >= 0.5 then
Update current search whales through Equation (5)
End if
End for
t = t + 1
End while
Return X*

C. Silhouette Coefficients

Silhouette coefficients mean a technique used to measure how well the object data are in clusters by providing brief graphical representation [11]. This method is the combination of cohesion and separation [11] [12]. In this study, silhouette coefficients are used to calculate the fitness value that each particle has in WOA. The following are calculation stages of silhouette coefficients:

1. Calculate average distance of data, e.g. i and all other data in a cluster.

$$a(i) = \frac{1}{|A|}\sum j \in_{A,j\neq i} C\ d(i,j) \quad (15)$$

Notes:
j = Other data in Cluster A.
d(i,j) = Distance between Data *i* and *j*

2. Calculate average distance of data-*i* and all data in other clusters, and take the least value.

$$d(i,C) = \frac{1}{|A|}\sum j \in C\ d(i,j) \quad (16)$$

Notes:
d(i,C) = average distance between Document i and all objects in other clusters or C where A ≠C.

$$b(i) = \min C \neq A\ d(i,C) \quad (17)$$

3. Calculate silhouette coefficients by using Equation (18):

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (18)$$

Average *s(i)* of all data in a cluster indicates how close the data similarity is in a cluster and how precise the data have been grouped. Clustering results are viewed from the silhouette coefficients. The greater these coefficients are, the better the clustering results produced are [11] [12].

III. Experimental Results

All algorithms were programmed in python language with Jupyter Notebook in the tools of Anaconda and executed on Intel Core, i5-7200U CPU @2.50 Ghz, 4 GB and a computer performing Microsoft Windows 10, Home Single Language. There were 10 datasets used to evaluate performance of the proposed algorithm in comparison to the one of other algorithms such as Wine, Heart, Sonar, Diabetes, DNA, Waveform, Page-blocks, Ann-thyroid, Letter-Recognition, and Shuttle. Research data were taken from UCI Machine Learning Repository (https://archive.ics.uci.edu/ml/machine-learning-databases/). Information on number of features, classes, and data can be seen in Table I. In this study, 80% of data were used for training and testing. The training data were implemented to see performance of WOA and other clustering algorithms.

TABLE I. CHARACTERISTICS OF INPUT DATASETS

Datasets	Number of Records	Number of Attributes	Number of Data (n)	
			Training	Testing
Wine	178	13	142	36
Heart	270	13	216	54
Sonar	208	60	166	42
Diabetes	768	8	614	154
DNA	2000	180	1600	400
Waveform	5000	21	4000	1000
Page-blocks	5473	10	4378	1095
Ann-thyroid	7200	21	5760	1440
Letter-recognition	20000	16	16000	4000
Shuttle	43500	9	34800	8700

Whale Optimization Algorithm Clustering (WOAC) was applied to a number of datasets using a parameter with early initialized clusters (K) = 10, number of iterations = 100, number of whales = 10, spiral value (b) = 0.618, decrease rate for a (da) = 2/mi, and vector r (random vector value between [0,1]). Next, it was compared to ACOC using a parameter with early initialized clusters (K) = 10, number of iterations = 100, pheromone exponent (α) = 1, heuristic exponent (β) = 2, number of ants = 10, pheromone evaporation factor (ρ) = 0.1, and initial pheromone value (τ) = 0.1. While Forgy applied such the parameter: number of clusters (K) = 10 and maximal number of iterations = 100, ISODATA included number of clusters (K) = 10, minimal threshold = 2, threshold of standard deviation = 1, threshold of minimal distance = 0.5, maximal threshold = 1, and maximal iterations = 100. Results obtained can be seen in Table II.

TABLE II. CLUSTERING EXAMINATION RESULTS OF DATASETS

Datasets	Clustering Algorithms	Time	Variance Ratios	SSE	Silhouette Coefficients	Formed Clusters
Wine	Forgy	00m.01s	0.6551	0.1382	0.4143	10
	ISODATA	00m.01s	0.1274	0.1547	0.4207	3
	ACOC	00m.14s	0.4438	0.0653	0.5446	2
	WOAC	00m.12s	0.4423	0.0502	0.5707	3
Heart	Forgy	00m.01s	0.2264	0.2708	0.5115	10
	ISODATA	00m.02s	0.2540	0.2480	0.5261	2
	ACOC	00m.22s	0.9560	0.0731	0.5273	3
	WOAC	00m.12s	0.8932	0.0912	0.6964	3
Sonar	Forgy	00m.06s	0.2588	0.6245	0.4632	10
	ISODATA	00m.09s	0.2845	0.6984	0.5104	10
	ACOC	01m.03s	0.1096	0.0294	0.6687	10
	WOAC	00m.46s	0.1033	0.1298	0.6902	10
Diabetes	Forgy	00m.12s	0.0174	0.5995	0.4239	10
	ISODATA	00m.15s	0.0174	0.5992	0.4913	7
	ACOC	03m.15s	0.5258	0.0698	0.5649	3
	WOAC	01m.12s	0.1122	0.0527	0.6235	3
DNA	Forgy	05m.19s	0.5109	1.2552	0.5122	10
	ISODATA	05m.25s	0.5156	1.2503	0.5275	3
	ACOC	34m.34s	0.6036	0.1594	0.5877	3
	WOAC	15m.36s	0.5287	0.1573	0.5975	3
Waveform	Forgy	05m.03s	0.1404	0.7647	0.5264	10
	ISODATA	05m.29s	0.1504	0.9477	0.5625	2
	ACOC	13m.21s	0.3827	0.1051	0.5828	9
	WOAC	06m.23s	0.2378	0.1093	0.6249	7
Page-blocks	Forgy	05m.06s	0.5352	0.5427	0.4204	10
	ISODATA	04m.13s	0.5995	0.5820	0.4132	7
	ACOC	12m.41s	0.7985	0.0819	0.4339	10
	WOAC	05m.32s	0.6325	0.0432	0.6204	10
Ann-thyroid	Forgy	05m.12s	0.3133	0.6526	0.4166	10
	ISODATA	03m.55s	0.3133	0.9151	0.4276	10
	ACOC	14m.58s	0.5828	0.0976	0.5266	10
	WOAC	07m.43s	0.3362	0.0762	0.66	10
Letter-recognition	Forgy	06m.01s	0.2172	0.5153	0.5205	10
	ISODATA	05m.50s	0.1932	0.5022	0.4539	10
	ACOC	32m.27s	0.2312	0.1560	0.6659	10
	WOAC	10m.26s	0.1892	0.1268	0.6718	10
Shuttle	Forgy	05m.42s	0.2309	0.6780	0.4758	10
	ISODATA	05m.47s	0.2308	0.6894	0.5184	9
	ACOC	38m.28s	0.2435	0.1758	0.5618	10
	WOAC	11m.12s	0.2273	0.1783	0.5761	10

Notes: m = minutes and s = seconds

Based on data provided in Table II, it is visible that clustering using Forgy always produces fixed number of clusters, whereas the ones using ISODATA, ACOC, and WOAC produce indefinite number of clusters. Table II further shows the comparison graphic of clustering time of the four methods. Figure 5, nevertheless, indicates the one of clustering duration of several datasets. More specifically, ACOC tends to require longer time to conduct clustering if compared to other methods.

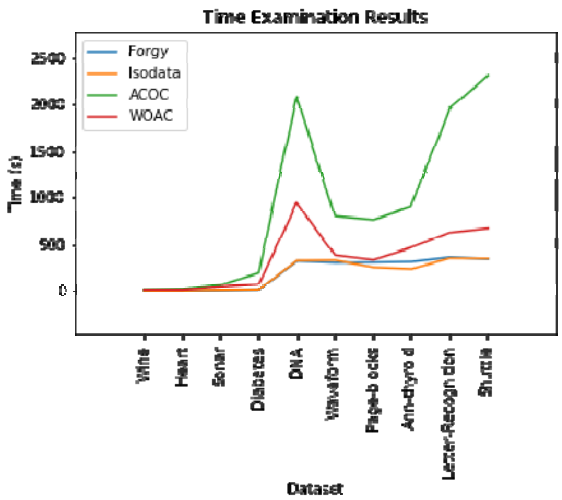

Fig 5. Comparison graphic of clustering time

In addition, referred to Table II, the comparison graphic of an ideal variance value can be drawn (see Figure 6). It is found that the four algorithms possess nearly identical variance values.

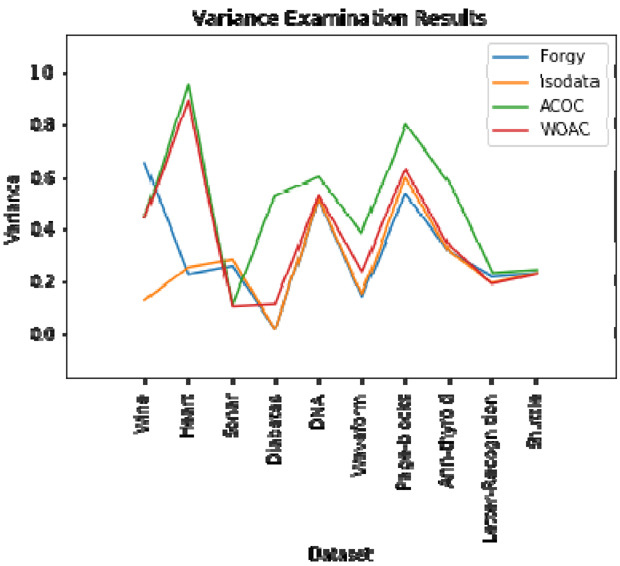

Fig 6. Comparison graphic of variance values

SSE was further displayed in the form of the graphic (see Figure 7). Referring to such the graphic, it is clear that WOA has less SSE. In other words, its clustering results are good.

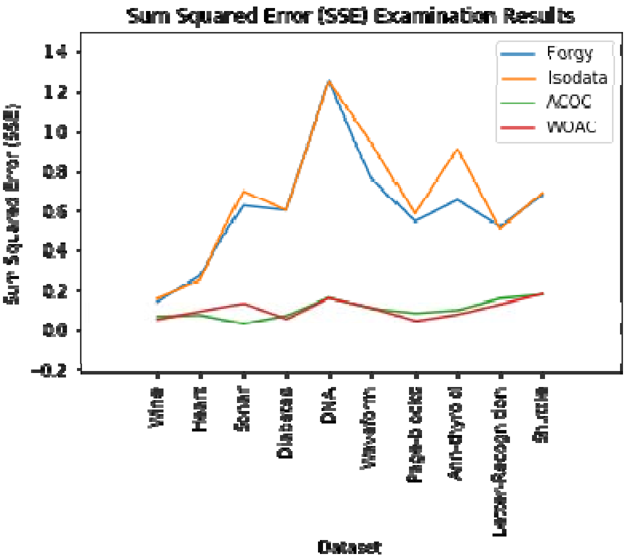

Fig 7. Comparison graphic of SSE

Another graphic was added to indicate comparison of silhouette coefficients of the four methods (see Figure 8). It is obvious that WOA has a high silhouette coefficient and represents good clusters. This is in accordance with the concept of such the coefficient that the greater it is, the better the clustering results are produced [11][12].

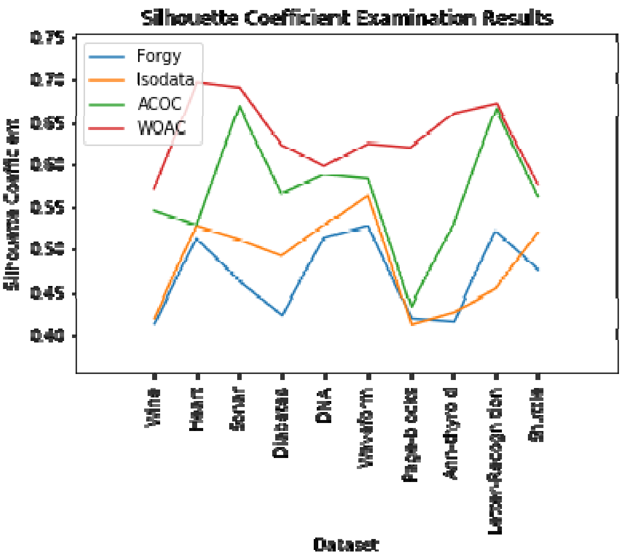

Fig 8. Comparison graphic of silhouette coefficients

Conclusion

Based on a series of conducted trials, it can be concluded that the WOA for clustering can solve problems of data clustering with good performance. Proven through silhouette coefficients and SSE obtained from a number of test datasets, clustering using WOA produces clusters which are better than those using ACO, Isodata, and Forgy. Overall, clustering results of WOA and the other three methods almost have the same variance values, and produce high intraclass similarity and low interclass similarity. These are based on distance between data vectors and the cluster center produced. To suggest future research, WOA can be developed and modified. It can also be combined with the use of other methods so that more optimal solutions with shorter processing time can be realized.

References

[1] M. Song, C.W. Günther, & W.M. Van der Aalst, "Trace clustering in process mining." In International Conference on Business Process Management (pp. 109-120). Springer, Berlin, Heidelberg. September, 2008.

[2] H. Jiang, S. Yi, J. Li, F. Yang, & X. Hu. "Ant clustering algorithm with K-harmonic means clustering." Expert Systems with Applications, 37(12), 8679-8684. 2010.

[3] H. Wang, W. Wang, J. Yang, P.S. Yu. "Clustering by pattern similarity in large data sets." In Proceedings of the 2002 ACM SIGMOD international conference on Management of data (pp. 394-405). ACM. June, 2002.

[4] S.J. Nanda, & G. Panda, G. "A survey on nature inspired metaheuristic algorithms for partitional clustering." Swarm and Evolutionary computation, 16, 1-18. 2014.

[5] C. Blum, A. Roli, A & M. Sampels. "Hybrid metaheuristics: an emerging approach to optimization". (Vol. 114). Springer. 2008.

[6] S. Mirjalili, & A. Lewis. "The whale optimization algorithm". Advances in engineering software, 95, 51-67. 2016.

[7] A.N. Jadhav & N. Gomathi. "WGC: Hybridization of exponential grey wolf optimizer with whale optimization for data clustering". Alexandria engineering journal, 57(3), 1569-1584. 2018.

[8] J. Nasiri & F.M. Khiyabani. "A whale optimization algorithm (WOA) approach for clustering". Cogent Mathematics & Statistics, 5(1), 1483565. 2018.

[9] P.J. Clapham. "The humpback whale. Cetacean Societies, field studies of dolphins and whales". Chicago: The University of Chicago, 173-196. 2000.

[10] D. Wiley, C. Ware, A. Bocconcelli, D. Cholewiak, A. Friedlaender, M. Thompson, & M. Weinrich. "Underwater components of humpback whale bubble-net feeding behaviour". Behaviour, 575-602. 2011.

[11] S. Aranganayagi & K. Thangavel. "Clustering categorical data using silhouette coefficient as a relocating measure". In International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007) (Vol. 2, pp. 13-17). IEEE. December, 2007.

[12] R. Lletı, M.C. Ortiz, L.A. Sarabia, & M.S. Sánchez. "Selecting variables for k-means cluster analysis by using a genetic algorithm that optimises the silhouettes". Analytica Chimica Acta, 515(1), 87-100. 2004.

[13] M.A. El Aziz, A.A. Ewees, A.E. Hassanien, M. Mudhsh, & S. Xiong. "Multi-objective whale optimization algorithm for multilevel thresholding segmentation". In Advances in soft computing and machine learning in image processing (pp. 23-39). Springer, Cham. 2018.