PalArch's Journal of Archaeology of Egypt / Egyptology

META-HEURISTIC ALGORITHMS FOR K-MEANS CLUSTERING: A REVIEW

Alan Fuad Jahwar
Akre Technical College,
Duhok Polytechnic University,
Duhok, Kurdistan Region, Iraq
alan.jahwar@dpu.edu.krd

Adnan Mohsin Abdulazeez
Duhok Polytechnic University
Duhok, Kurdistan Region, Iraq
adnan.mohsin@dpu.edu.krd

Alan Fuad Jahwar, Adnan Mohsin Abdulazeez. Meta-Heuristic Algorithms for K-means Clustering: A Review-Palarch's Journal of Archaeology of Egypt/Egyptology 17(7), ISSN 1567-214x

ABSTRACT

The increase in the data available attracted the concern of clustering approaches to integrate them coherently and to identify patterns for big data. Hence, Meta-Heuristic algorithms can be better than standard optimization algorithms in some instances. Previously, optimization issues have been considered as significant weaknesses in the K-means algorithm is one of the simplest methods for clustering. and with less additional information it can easily solve the optimization problem. In this paper, a review of clustering k-means algorithm and meta-heuristics algorithms are reviewed.

Keyword: Machine Learning, Clustering, K-Means, Meta-Heuristic algorithms.

1 INTRODUCTION

The concepts of Machine Learning (ML) come from the domains of computer science and Artificial intelligence (AL),ML deals with systems that can learn from data instead of only executing the programmed commands overtly [1,2,3]. Furthermore, ML is closely linked to optimization and statistics, which brought their theories and approaches to the field. ML is utilized in different computing missions where constructing and programming rule-based, overt algorithms is not feasible. In certain cases, ML, pattern recognition, and data mining share their background [4,5,6].

Samuel (1959) defined ML as the domain that offered computers the capability of learning without being obviously programmed [7,8]. Broadly speaking, the boundary between ML and data mining is sometimes blurs because both adopt similar approaches, and they are considerably interrelated. However, their difference is that ML is dedicated to prediction centered on known properties which can be taught from training data [9]. Meanwhile, data mining is devoted to the discovery of the (previously) unknown features in the data. These two areas share many characteristics: data mining uses plentiful ML approaches with a various goal in mind. On the other hand, ML uses data mining approaches as unsupervised or supervised learning or as a stage that comes before the enhancement of learning accuracy. The current study: however, is ML centered.

Concerning its functions, there are three kinds of ML, which are the supervised, Semi-supervised and Unsupervised learning. In supervised learning, the teacher provides example inputs and desired outputs, in order to learn a universal rule for mapping inputs to outputs. In supervised learning, the ML task is inferring from tagged training data [10]. Classification is among the instances of the supervised learning that uses labeled data to solve certain problems. In the supervised learning, the ML seeks to trace functions in the labeled training data [11,12,13].

It consists of a group of training examples and every example represents a pair that is made up of an input (usually a vector), and an anticipated output value (supervisory signal). Algorithms that harness Supervised learning analyses training data and creates a deduced function capable of mapping new cases. The ideal setting is expected to help an algorithm select properly the class labeling for unsupervised examples. For this objective, the learning algorithm has to perform generalization from training data to unsupervised cases "reasonably"[14]. The functioning of unsupervised learning algorithms is determined by unlabeled instances, i.e., input where the desired output is unknown. In such a case, the aim is the discovery of the data structure, for instance by means of a cluster analysis, starting from inputs to outputs to generalize a map [15].

K-means has a high clustering speed and performs well in large data sets, but it has poor clustering accuracy, is vulnerable to noise and isolated data, and the value of K needs to be calculated in advance. In order to address the weaknesses of the K-means algorithm, scientists proposed changes in various angles [16,17]. Intelligent algorithms with high global optimization abilities are commonly used in modern industries [18]. Since there are a need for efficient and robust computational algorithms that can solve optimization problems in different fields; this is the practical utility of optimization. [19,20]. Finding a best solution to a problem is optimization. An optimization problem is defined as minimizing or maximizing some function. The optimization problems focus on three factors: (1) a target objective function to minimize or maximize. (2) A set of variables influencing the objective function. (3) A set of constraints that make some unknowns equivalent but exclude others [21].

In many optimization problems there are more than one local optimum solutions. Therefore, it is crucial to choose an appropriate an optimization method that will not look in the neighborhood of the best solution, misleading the search process. Causing it to get stuck in local minima. Besides, it should also have a mechanism to balance between local and global search. Optimization problems of both the mathematical and combinatorial types are solved using several methods. If the optimization problem is difficult to solve or the search space is big, then classical mathematics would be incapable of finding an optimal solution [22,23].

2 Meta-Heuristic Algorithms

A Meta-heuristic is a continuous iteration of different concepts and structures to find near-optimal solutions. Meta-heuristic algorithms are among these approaches to improve complex problems [24,25]. Meta-heuristic algorithms usually find the global optimum more quickly than ordinary stochastic algorithms. The algorithms of meta-heuristics consist of intensification and diversification or (exploitation and exploration) [26]. Most meta-heuristic algorithms are nature-inspired and include Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO) [27].

2.1 Meta-heuristic Method for Clustering

Heuristic is a method that is faster than the classical method and can get a good approximation in many cases. By optimality, accuracy, completeness, or precision, an improvement may be achieved by more rapid trading. Heuristic thoughts and easy solutions are a shortcut for solving complex problems [28]. A meta-heuristic is a higher-level heuristic that is used to find, generate, or select a lower-level heuristic or procedure (partial search algorithm). and which might suggest a solution to an optimization problem, in particular with imperfect information and limited computational capacity [29].

Meta-heuristics often do not make any assumptions about the optimization problem being solved. As a result, they can be applied to a wide range of problems, compared to iterative methods and optimization algorithms, meta-heuristic is not capable of providing optimal solutions for all problem classes [30]. By implementing heuristic search methods in some ways, the found solution depends on the collection of random variables generated [31]. Meta-heuristic is able to search over a large set of feasible solutions with less computational efforts than iterative methods, algorithms, or simple heuristics. The k-means algorithm is improved by a meta-heuristic method [29].

3 Swarm Intelligence (SI)

The best pattern discovery technology is based on optimization and important in information discovery and data mining (KDD). Cluster analysis is known as a quick and easy way to analyze complex data set. These datasets are, in a way, very complex, and so there is an opportunity for clustering techniques to be applied. Different and various optimization techniques were used to examine clustering solutions. Swarm intelligence (SI) is one of the optimization approaches that has achieved tremendous success in different disciplines [32].

3.1 Computational Swarm Intelligence

Swarm Intelligence (SI) is an efficient computation paradigm suited for adaptive systems. This adds to genetic adaptation and social observation when discussing with the application of SI. In the literature, SI involves the installation of collective intelligence of groups of simple agents, which is applied to solve problem-solving tools such as the school's fish, bird flocks and insect colonies (like ants, termites and honeybees) by performing collective activities. In the 1980s, ethnologists conducted several studies in which they modeled the swarm behavior and concluded interesting observations. Each of the individuals within the swarm possesses a stochastic behavior in reaction to the perception of the environment. Local rules which are independent from the global rules and interactions between the self-organized agents lead to the emergence of collective intelligence. There is self-organization within swarms as the interactions on the local level lead to a global level response[33,34]. These trajectory tracking algorithms show how a decentralized, self-organized pattern can emerge in the collective foraging behavior of animals [35]. The major principles which express that swarm intelligence is an intelligent behavior are:

- The swarm is capable of both processing spatial and temporal data (the proximity principle).
- The swarm should be able to adapt to changing conditions, such as food quality and location (the principle of quality).
- The swarm should not provide all its resources to a narrow range o f nodes, and should allocate the same resources to all the nodes in t he swarm (the principle of diverse response).
- The swarm does not change its behavior if the environment is fluct uating (the stability principle).
- The swarm should have the capability to change itself whenever ap propriate (the principle of adaptability) [36].

3.2 Particle Swarm Optimization (PSO)

PSO is both feasible and has successfully been implemented on other nonlinear problems like network training and fuzzy control [37]. When it comes to PSO, the elements form a collective swarm in finding a solution for the objectives. Every particle has two attributes, including the position and the velocity, which can be used to determine its future direction [38]. The population must continue the iteration process until the target is satisfied [39]. PSO also known as the bird's algorithm, as well as other meta-heuristic algorithms, allow you to construct a random population of individuals and decide the optimal value for each individual. If a particle enters a new location from its previous location, it can be moved to the Personal Best, or the Global Best, or it can be moved directly to its previous location. [40].

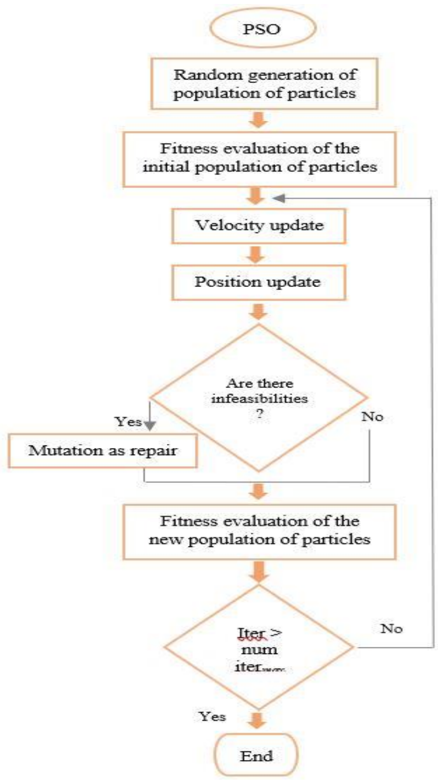

Figure 1: Flowchart PSO [41]

3.3 Artificial Bee Colony (ABC) Algorithm

The ABC algorithm was developed by Karaboga inspired by the activity of honeybee swarms. In the ABC colony, the bees can be arranged in three groups: workers, onlookers, and scouts. The half of the colony is mainly composed of workers, and the other half is made up of spectators. The actively employed bees are considered responsible for searching and analyzing food sources. They transfer their food to other insects such as bees. The onlookers pick the best food sources out of the foods that bees found in the first place. Once the quality of the natural food supply drops significantly, the bees will quit for a new source of food [42,43].

3.4 Ant Colony Optimization (ACO)

ACO is based on the natural behavior of ant-colony and individual workers. When ants are searching for food, it means they instinctively seem to determine the most optimum route to obtain food. This observed behavior is the basis for ACO. Imagine two ants walking for food down different ways to find food supply. When ants walk, they release airborne chemicals that cause decay over time. The ant which starts the trip in the direction of shorter route will do so quicker than the other ant; thus, reinforcing the pheromone trail. These other ants will understand this signal and be influenced to follow the direction. According to [44], the ACO optimization algorithm has three major steps that constitute the core of the optimization phase.

1. Constructor ants. This is the algorithm which "ants" use to accomplish some sort of development or improvement by way of self-organization or growth.
2. Pheromone evaporates. This is the process in which pheromones are reduced by using "local" information for certain solutions; this step is also often called a local update. This step ensures that the ACO does not converge prematurely to a single solution.
3. Daemon Actions. This one is known as decisions involving global information on the problem of optimization. Note the difference between local and global in this example. Like step 2, step 3 is also referred to as global update.

4 Clustering Algorithms

Data clustering is a study area that is widely approached in data mining and ML fields because it is applied to segmentation, summarization, learning, information retrieval, pattern recognition and target marketing, data mining[25,46,47,48], and text mining [49,50]. The clustering algorithms are divided into two categories, including hierarchical clustering and partitional clustering.[51]. Clustering algorithms group sets of objects so that objects in a given set or cluster share more similarity than the objects of other clusters. Clustering is primarily executed in exploratory data mining and this strategy is utilized usually for analyzing statistical data, and it is exploited in a number of domains like ML, retrieving the information, pattern recognizing, bioinformation, and analyzing image [14,52]. Cluster analysis is a general task to be fulfilled. This task may be accomplished by a variety of algorithms which differ significantly in terms of the notion which is relevant to constituting a cluster and the method of discovering them. The cluster model typically involves short distances between the cluster components, condensed areas within data space and certain statistical distributions [53,54].

Clustering analysis is the most predominant tool in microarray data analysis, that can group genes with similar expression patterns but under different experimental settings or taken from other tissues because genes with analogous expression profiles commonly function similarly, genes that have unknown functions are predictable based on the resulting class [55]. Clustering is exploited in many fields, for instance retrieving information [56,57]. Clustering are helpful in finding more rapidly relevant information [58]. It helps scholars to be updated with the newest discoveries in their research areas. In the recent period, clustering has drawn the attention of numerous scholars due to its applications in classifying, decision making, extracting information, and analyzing patterns [59]. However, the shared starting point is that they are groups of data objects. As found by different algorithms, cluster concept considerably shows a discrepancy in its features. Understanding such "cluster models" leads to understanding variations found among various algorithms [60,61]. For each cluster, the single-sample approaches provide an example and assign data points that minimize the number of distances between data points and their closest examples. A common method for k-means is k-medoids clustering, which is very similar to k-means [62]. Figure 1 Differentiates between the k-mean and k-medoid of a cluster.

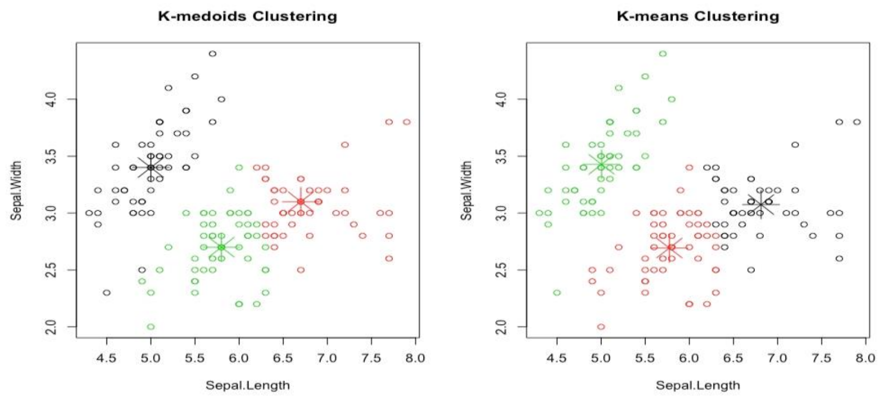

Figure 2: Differentiates between the k-mean and k-medoid of a cluster [62]

K-means Vs K-medoids

- Both require K to be specified in the input.
- K-medoids eliminates outliers in the data.
- K-medoids is more expensive to perform.
- Both methods assign each instance exactly to one cluster.

Fuzzy partitioning techniques relax this

- Partitioning methods generally are good at finding spherical shaped clusters.
- They are suitable for small and medium sized datasets.

Extensions are required for working on large data sets [62].

4.1 K-Means

K-means clustering, commonly used for data mining analysis, is a method of vector quantizing originating from signal treatment. The objective of K-means is to divider n observations into K clusters; all observations are part of the cluster that works as a prototype of a Cluster[63].The K-Means Approach is commonly used and is an iterative process that begins with the initial partitioning and then converges on the best results with decreasing the sum squared error (SSE) [64,65].The problem has proved to be an NP-hard problem. While there are several effective heuristic algorithms that are able to quickly find the algorithm is equivalent to the expectation-maximization algorithm because of the methods evaluated of how they move towards a global optimum through an iterative refinement method [66].

The objective of the classical K-means clustering method is find the set C of K clusters Cj with cluster mean cj for the sake of decreasing the amount of the squared errors [17]. As show in Equation 1.

$$E = \sum_{i}^{k} = 1 \sum xi \in cj \| cj - xi \|^2 \quad (1)$$

E is the addition of (SSE) of objects having cluster means for K cluster, || … || refers to distance Mertie between a cluster mean and a data point xi Cj. As show in Equation 2.

$$\|\boldsymbol{x} - \boldsymbol{y}\| = \sqrt{\sum_{i=1}^{v} |\mathrm{xi} - \mathrm{yi}|^2} \quad (2)$$

A flowchart of K-means clustering has been illustrated which is made up of six essential stages. First, preliminary value of centroids: Let ($C1$, $C2$, ...) represents centroids harmonize. Second, the objects' distance is calculated using the cluster centroid and the objects in the cluster. The distance Euclidean is used and the distance matrix is then calculated with the iteration 0. Each column in the matrix of distance means an object. The distance of matrix in the first row matches the distance of every object to the 2nd row and the first centroid stands for the distance of every object in the 2nd centroid. At 3rdrow, clustering of objects: Allocate every object on the basis of least distance. At 4throw iteration-1, determining the centroids: by identifying the components of all groups, the new centroid of every set and it is computed on the basis of these memberships which are new. At fifth row, repeating from step 2. At sixth row, the last iteration grouping is compared and this iteration states that groups are not moved by the objects So, K- means clustering computation means that it has become stable and there is no need of iteration anymore [17].

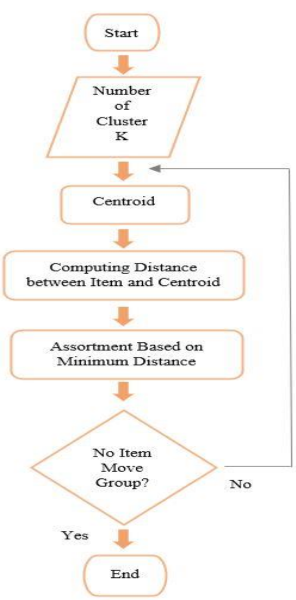

Figure 3: Flowchart of K-means Clustering [17]

4.2 Related Work of K-Means

Despite the fact that MacQueen used the term "K-means" for the first time in 1967 [66] the idea dates back to Steinhaus in 1957 [67]. The pulse-coding modulation standard algorithm has been proposed for the first time by Lloyd in 1957; though, it was not published until1982 [68]. Moreover, Forgy published the same approach for its similarity in 1965, and this why it is sometimes named after him [69]. While Hartigan (1979) also published a more efficient version. Standard K-means algorithms employ the strategy of iterative refinement. Due to its pervasiveness, it is called K-means algorithm and also Lloyd's algorithm, particularly in computer sciences. At the beginning, the algorithm is offered a set of K-means $m1$,mk, then it continues by alternating between two stages [70]. The first stage is allocating where every observation is allocated to a cluster that its mean gives up the minimum within-cluster sum of squares (WCSS). As sum of squares is the squared Euclidean distance, it is called "nearest" mean [71].

The second stage is the updating where new means are estimated to stand for the centroids of the observations within the new clusters. When the allocating no longer changes, the algorithm can be called converged. Since both stages optimizes WCSS objective and there is only a finite number of partitioning, the algorithm is supposed to be altered to a (local) optimum. By using this algorithm, the global optimum cannot be guaranteed to be found [72]. Initialization approaches utilized in K-means algorithm are Forgy and Random Partition [69]. Believe that typically the random partition approach is favorable for instance, regarding K-harmonic means and fuzzy K-means. Forgy's approach of initialization for standard K-means and expectation maximization algorithms is superior to existing approaches.[73].

Since it is an exploratory algorithm, its convergence to the global optimum is not assured; and the outcomes may rely on the preliminary clusters [74]. Since this algorithm is typically so quick, it is usually executed more than one time with different beginning conditions. However, in its very bad cases, K-means so slowly converges [75]. One concern with the K-means clustering algorithm is that it could result in large intra-cluster distances. Therefore, the distance between members of cluster and the cluster center is long and members of cluster are different from each other. To solve this problem the optimization algorithm can be used. A few of the optimization methods is PSO, that provides better results. In addition, to this problem. A hybrid of PSO and K-means clustering algorithm can be used to solve the problem, and any problems can be eliminated [76].

Table 1 shows which methods have been proposed in the literature to increase the accuracy of k-means clustering algorithm.

Table 1: Meta-Heuristic Review for K-means Clustering

Ref.	Year	Methods	Problems	Datasets	Results
[77]	2020	○ PSO ○ PSOFKM ○ PSOLF-KHM ○ PSOM	Initiating the cluster centres, catch up with local point of optimum.	○ IRIS ○ WINE ○ GLASS ○ HEART ○ CANCER ○ ECOLI ○ CREDIT ○ YEAST	PSOM is able to solve the best compared to PSO, PSOFKM and PSOLFKHM for all datasets.
[78]	2020	○ A novel (Uk-means)	A variety of K-means and its extensions are always affected by initializations and a necessary number of clusters a priori.	○ Synthetic Datasets ○ UCI Datasets ○ Medical Datasets ○ Image Datasets	The actual results indicate that the U-k-means clustering algorithm is more efficient.
[79]	2020	○ ABC ○ ACO ○ BBO ○ CA ○ DE ○ GA ○ HS ○ IWO ○ PSO ○ TLBO	Making the most suitable cluster centroids is a key goal for a successful k-means clustering operation.	○ synthetic dataset ○ real-world datasets	A comparison is performed on various real-world datasets.
[80]	2020	○ PSO ○ K-means	Image is de-noised and colorized.	○ Lena ○ Tree ○ Flower	The algorithm can segment images with higher accuracy and higher efficiency than (PSOK).
[81]	2019	○ SK-means ○ EM ○ PSO ○ SCPSO	The problem of internal inconsistency in group of documents in a variety of areas.	○ Reuters ○ 20Newsgroup ○ TDT2	The proposed SCPSO algorithm is better than other techniques for clustering.
[82]	2019	Combined PSO with K-means (PSOKM)	Clustering analysis groups, the text into similar clusters, and the text in different clusters are the most dissimilar.	○ SSE ○ XB ○ DB	PSOKM does better than the others.
[83]	2019	○ CLARA ○ K-Means	Comparing two techniques, (CLARA) clustering and K-Means clustering.	○ Iris	CLARA clustering is better than the K-Means.

[84]	201	○ AC Clustering ○ GA Clustering ○ Firefly Algorithm ○ Firefly Clustering	NP hard problem: find minimum tour length and return to starting node.	○ eil51 ○ eil76 ○ pr76 ○ ulysses16	Prove the effectiveness of k-means clustering in solving the TSP with the k-means algorithm.
[85]	2018	○ GA ○ PSO ○ GODLIKE	Main characteristic or similarity regions of an image segmentation.	○ Mountain ○ Pepper ○ Lena ○ Boat ○ Cameraman ○ Brain ○ Outdoor ○ Building A ○ Building B	The simulation result showed that the proposed method provides better output.
[86]	2017	○ PS-BCO-K ○ K-PS-BCO	Data clustering	○ Iris ○ Wine ○ Cancer ○ CMC ○ HV	The proposed algorithms give a more accurate quality solution than some well-known heuristic algorithms.

5 Discussion

From the literature that has been done above it is shown that the metaheuristic algorithms are proved efficient at finding the local optima problem of the group clustering. However, because these algorithms are implemented empirically and their effectiveness is proved by the research, it is not up to challenge their efficiency.A review and study on the K-means algorithm found some shortcomings. The main objective of the above table is to review the K-means algorithm and meta-heuristic algorithms to improve the K-means algorithm. The main focus is on the clustering algorithm review. The metrics in the studies are reviewed. The databases that were used are investigated. One of the ways to reduce the shortcomings of the K-means clustering is by using hybrid method.

Based on the review that has been done earlier, several authors suggested methods to address the shortcomings of partitional clustering issue without unique data and sample such that optimum number of clusters can be computed. Metaheuristic approaches proved that it could get rid of such problems as aforementioned. This is since metaheuristic algorithms can detect and use the local optima and can provide an efficient solution for finding number of clusters in the specified dataset. Based on the table it has been shown that the PSO method has been used in many studies to overcome the shortcomings of clustering methods. The study of [79] has used several optimization techniques to improve the k-means clustering, based on the evaluation results it has been shown that the PSO method has got higher results compared to other techniques such as ABC, ACO, GA, and so on. The study of [80] used k-means method in segmentation task, then combined PSO with k-means called (PSOK) which got higher results. More so, based on the study [84], it has been proven that the firefly technique also can improve the k-means, where they used ACO, GA, Firefly, and each of them with k-means in TSP solution, it is shown that firefly can improve K-means better that other techniques.

6 Conclusion

A literature study on the efficacy of K-means showed that this technique has some shortcomings. As K-means is a popular clustering algorithm due to its simplicity and efficiency, there is a need for its improvement. The K-means clustering algorithm can be applied with meta-heuristic algorithms to enhance the performance, Since the clustering algorithm is an unsupervised algorithm, it is more efficient in gathering information than algorithms used for improved K-means which have extra information necessary to solve problem.

Acknowledgement The authors would like to acknowledge Duhok Polytechnic University for providing all financial support and support for this study.

References

[1] Ackerman, M. S. (2000). The intellectual challenge of CSCW: the gap betweensocial requirements and technical feasibility. *Human–Computer Interaction*15(2-3): 179-203.

[2] Abdulqader, D. M., Abdulazeez, A. M., &Zeebaree, D. Q. (2020). Machine Learning Supervised Algorithms of Gene Selection: A Review. *Machine Learning*, *62*(03).

[3] Sulaiman, M. A. (2020). Evaluating Data Mining Classification Methods Performance in Internet of Things Applications. Journal of Soft Computing and Data Mining, 1(2), 11-25.

[4] Chakrabarti, S. (2003). Mining the Web: Discovering knowledge from hypertext data. Morgan Kaufmann.

[5] Zebari, D. A., Zeebaree, D. Q., Saeed, J. N., Zebari, N. A., & Adel, A. Z. (2020). Image Steganography Based on Swarm Intelligence Algorithms: A Survey. people, 7(8), 9.

[6] Zeebaree, D. Q., Haron, H., Abdulazeez, A. M., &Zebari, D. A. (2019, April). Machine learning and Region Growing for Breast Cancer Segmentation. In 2019 International Conference on Advanced Science and Engineering (ICOASE) (pp. 88-93). IEEE.

[7] Simon, P. (2013). Too Big to Ignore: The Business Case for Big Data. John Wiley & Sons.

[8] Bargarai, F., Abdulazeez, A., Tiryaki, V., &Zeebaree, D. (2020). Management of Wireless Communication Systems Using Artificial Intelligence-Based Software Defined Radio.

[9] Huang, T.-M., V. Kecman and I. Kopriva (2006). Kernel based algorithms for mining huge data sets: Supervised, semi-supervised, and unsupervised learning. Springer

[10] Glickman, M., J. Balthrop and S. Forrest (2005). A machine learning evaluation of an artificial immune system. Evolutionary Computation 13(2): 179-212.

[11] Mohri, M., A. Rostamizadeh and A. Talwalkar (2012). Foundations of machine learning. MIT press.

[12] Sulaiman, D. M., Abdulazeez, A. M., Haron, H., & Sadiq, S. S. (2019, April). Unsupervised Learning Approach-Based New Optimization K-Means Clustering for Finger Vein Image Localization. In *2019 International Conference on Advanced Science and Engineering (ICOASE)* (pp. 82-87). IEEE.

[13] Mohammed, N. N., &Abdulazeez, A. M. (2017, June). Evaluation of partitioning around medoids algorithm with various distances on microarray data. In *2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)* (pp. 1011-1016). IEEE.

[14] Jain, A. K. (2010). Data clustering: 50 years beyond K-means. Pattern Recognition Letters 31(8): 651-666.

[15] Ayodele, T. O. (2010). Types of machine learning algorithms. New advances in machine learning, 3, 19-48.

[16] Rong, Y. (2020, June). Staged text clustering algorithm based on K-means and hierarchical agglomeration clustering. In *2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)* (pp. 124-127). IEEE.

[17] Zeebaree, D. Q., Haron, H., Abdulazeez, A. M., &Zeebaree, S. R. (2017). Combination of K-means clustering with Genetic Algorithm: A review. *International Journal of Applied Engineering Research*, *12*(24), 14238-14245.

[18] Zhang, L. (2018, July). A Gravitational Artificial Bee Colony Optimization Algorithm and Application. In *2018 Eighth International Conference on Instrumentation & Measurement, Computer, Communication and Control (IMCCC)* (pp. 1839-1842). IEEE.

[19] Alazzam, A., & Lewis, H. W. (2013). A new optimization algorithm for combinatorial problems. *IJARAI) International Journal of Advanced Research in Artificial Intelligence*, *2*(5).

[20] Zebari, N. A., Zebari, D. A., Zeebaree, D. Q., & Saeed, J. N. Significant features for steganography techniques using deoxyribonucleic acid: a review.

[21] Said, G. A. E. N. A., Mahmoud, A. M., & El-Horbaty, E. S. M. (2014). A comparative study of meta-heuristic algorithms for solving quadratic assignment problem. *arXiv preprint arXiv:1407.4863*.

[22] Baghel, M., Agrawal, S., &Silakari, S. (2012). Survey of metaheuristic algorithms for combinatorial optimization. *International Journal of Computer Applications*, *58*(19).

[23] Zebari, D. A., Zeebaree, D. Q., Abdulazeez, A. M., Haron, H., & Hamed, H. N. A. (2020). Improved Threshold Based and Trainable Fully Automated Segmentation for Breast Cancer Boundary and Pectoral Muscle in Mammogram Images. *IEEE Access*, *8*, 203097-203116.

[24] John Silberholz and Bruce Golden, "Comparison of Meta-heuristic " Handbook of Meta-heuristic algorithms International Series in perations Research & Management Science Volume 146, pp 625-640,2010.

[25] Stegherr, H., Heider, M., &Hähner, J. (2020). Classifying Metaheuristics: Towards a unified multi-level classification system. *Natural Computing*, 1-17.

[26] Eesa, A. S., Brifcani, A. M. A., & Orman, Z. (2013). Cuttlefish algorithm-a novel bio-inspired optimization algorithm. *International Journal of Scientific & Engineering Research*, *4*(9), 1978-1986.

[27] Ruiz-Vanoye, J. A., Díaz-Parra, O., Cocón, F., Soto, A., Arias, M. D. L. Á. B., Verduzco-Reyes, G., & Alberto-Lira, R. (2012). Meta-heuristics algorithms based on the grouping of animals by social behavior for the

traveling salesman problem. *International Journal of Combinatorial Optimization Problems and Informatics*, *3*(3), 104-123.

[28] Mohtashami, A., M. Tavana, F. J. Santos-Arteaga and A. Fallahian-Najafabadi (2015). A novel multi-objective meta-heuristic model for solving crossdocking scheduling problems. Applied Soft Computing 31: 30-47.

[29] Blum, C., J. Puchinger, G. R. Raidl and A. Roli (2011). Hybrid metaheuristics in combinatorial optimization: A survey. Applied Soft Computing 11(6): 4135- 4151.

[30] Blum, C. and A. Roli (2003). Metaheuristics in combinatorial optimization:Overview and conceptual comparison. *ACM computing surveys (CSUR)*35(3): 268-308.

[31] Bianchi, L., M. Dorigo, L. M. Gambardella and W. J. Gutjahr (2009). A survey onmetaheuristics for stochastic combinatorial optimization. *Natural Computing:an international journal* 8(2): 239-287.

[32] Alam, S., Dobbie, G., Koh, Y. S., Riddle, P., & Rehman, S. U. (2014). Research on particle swarm optimization-based clustering: a systematic review of literature and techniques. Swarm and Evolutionary Computation, 17, 1-13.

[33] Kennedy, J. (1997). The particle swarm: social adaptation of knowledge. Evolutionary Computation, 1997., IEEE International Conference on, IEEE.

[34] Abdulazeez, A., Salim, B., Zeebaree, D., &Doghramachi, D. (2020). Comparison of VPN Protocols at Network Layer Focusing on Wire Guard Protocol.

[35] Hartigan, J. A. and M. A. Wong (1979). Algorithm AS 136: A k-means clustering algorithm. Applied statistics: 100-108.

[36] Eberhart, R. C., Shi, Y., & Kennedy, J. (2001). Swarm intelligence. Elsevier.

[37] Feng, Z. K., Niu, W. J., & Cheng, C. T. (2018). Optimization of hydropower reservoirs operation balancing generation benefit and ecological requirement with parallel multi-objective genetic algorithm. *Energy*, *153*, 706-718.

[38] Zebari, D. A., Haron, H., Zeebaree, D. Q., & Zain, A. M. (2019, August). A Simultaneous Approach for Compression and Encryption Techniques

Using Deoxyribonucleic Acid. In 2019 13th International Conference on Software, Knowledge, Information Management and Applications (SKIMA) (pp. 1-6). IEEE.

[39] Feng, Z. K., Niu, W. J., & Cheng, C. T. (2018). Optimizing electrical power production of hydropower system by uniform progressive optimality algorithm based on two-stage search mechanism and uniform design. *Journal of Cleaner Production*, *190*, 432-442.

[40] Akbarpour, A., Zeynali, M. J., &Tahroudi, M. N. (2020). Locating optimal position of pumping Wells in aquifer using meta-heuristic algorithms and finite element method. *Water Resources Management*, *34*(1), 21-34.

[41] Roch-Dupré, D., Gonsalves, T., Cucala, A. P., Pecharromán, R. R., López-López, Á. J., & Fernández-Cardador, A. Determining the optimum installation of energy storage systems in railway electrical infrastructures by means of swarm and evolutionary optimization algorithms. *International Journal of Electrical Power & Energy Systems*, *124*, 106295.

[42] Karaboga, D. (2005). *An idea based on honey bee swarm for numerical optimization* (Vol. 200, pp. 1-10). Technical report-tr06, Erciyes university, engineering faculty, computer engineering department.

[43] Karaboga, D., & Ozturk, C. (2011). A novel clustering approach: Artificial Bee Colony (ABC) algorithm. *Applied soft computing*, *11*(1), 652-657.

[44] Bianchi, L., Dorigo, M., Gambardella, L. M., &Gutjahr, W. J. (2009). A survey on metaheuristics for stochastic combinatorial optimization. *Natural Computing*, *8*(2), 239-287.

[45] Sahoo, G. (2017). A two-step artificial bee colony algorithm for clustering. *Neural Computing and Applications*, *28*(3), 537-551.

[46] Katari, V., S. C. Satapathy, J. Murthy and P. P. Reddy (2007). Hybridized improvedgenetic algorithm with variable length chromosome for image clustering.*IJCSNS International Journal of COmputer Science and Network Security*7(11): 121-131.

[47] Rani, M. S., & Babu, G. C. (2019). Efficient query clustering technique and context well-informed document clustering. In *Soft Computing and Signal Processing* (pp. 261-271). Springer, Singapore.

[48] Kumar, Y., & Sahoo, G. (2015). A hybrid data clustering approach based on improved cat swarm optimization and K-harmonic mean algorithm. *AI communications*, *28*(4), 751-764.

[49] Büyüksaatçı, S., &Baray, A. (2016). A brief review of metaheuristics for document or text clustering. In *Intelligent Techniques for Data Analysis in Diverse Settings* (pp. 252-264). IGI Global.

[50] M. Muhammad, D. Zeebaree, A. M. Abdulazeez, J. Saeed, and D. A. Zebari, "A Review on Region of Interest Segmentation Based on Clustering Techniques for Breast Cancer Ultrasound Images", JASTT, vol. 1, no. 3, pp. 78 - 91, Jun. 2020.

[51] Figueiredo, E., Macedo, M., Siqueira, H. V., Santana Jr, C. J., Gokhale, A., & Bastos-Filho, C. J. (2019). Swarm intelligence for clustering—A systematic review with new perspectives on data mining. *Engineering Applications of Artificial Intelligence*, *82*, 313-329.

[52] Sahoo, G. (2017). A two-step artificial bee colony algorithm for clustering. *Neural Computing and Applications*, *28*(3), 537-551.

[53] Jain, A. K. and S. Maheswari (2012). Survey of recent clustering techniques in datamining. *Int. J. Comput. Sci. Manage. Res* 1: 72-78.

[54] Adeen, I. M. N., Abdulazeez, A. M., &Zeebaree, D. Q. Systematic Review of Unsupervised Genomic Clustering Algorithms Techniques for High Dimensional Datasets.

[55] Yang, C.-S., L.-Y. Chuang and C.-H. Ke (2008). Comparative particle swarmoptimization (CPSO) for solving optimization problems. *Research,Innovation and Vision for the Future, 2008. RIVF 2008. IEEE InternationalConference on*, IEEE.

[56] Fayyad, U., G. Piatetsky-Shapiro and P. Smyth (1996). The KDD process forextracting useful knowledge from volumes of data. *Communications of theACM* 39(11): 27-34.

[57] Ganesh, A. D. S. H., D. P. Cindrella and A. J. Christy (2015). A REVIEW ONCLASSIFICATION TECHNIQUES OVER AGRICULTURAL DATA.

[58] Jiang, D., C. Tang and A. Zhang (2004). Cluster analysis for gene expression data: Asurvey. *Knowledge and Data Engineering, IEEE Transactions on* 16(11):1370-1386.

[59] Berkhin, P. (2006). A survey of clustering data mining techniques.*Groupingmultidimensional data* 25-71, Springer.

[60]Bezdek, J. C. and N. R. Pal (1998). Some new indexes of cluster validity. *Systems,Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on* 28(3):301-315.

[61] Warren Liao, T. (2005). Clustering of time series data—a survey. Pattern Recognition 38(11): 1857-1874.

[62] Sammut, C., & Webb, G. I. (2017). *Encyclopedia of machine learning and data mining*. Springer.

[63] Duwairi, R. and M. Abu-Rahmeh (2015). A novel approach for initializing the spherical K-means clustering algorithm. Simulation Modelling Practice and Theory 54: 49-63.

[64] Hussain, S. F., &Haris, M. (2019). A k-means based co-clustering (kCC) algorithm for sparse, high dimensional data. *Expert Systems with Applications*, *118*, 20-34.

[65] Zeebaree, D. Q., Haron, H., Abdulazeez, A. M., &Zebari, D. A. (2019, April). Trainable Model Based on New Uniform LBP Feature to Identify the Risk of the Breast Cancer. In *2019 International Conference on Advanced Science and Engineering (ICOASE)* (pp. 106-111). IEEE.

[66] MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, California, USA.

[67] Gayathri, R., A. Cauveri, R. Kanagapriya, V. Nivetha, P. Tamizhselvi and K. P.Kumar (2015). A Novel Approach for Clustering Based on Bayesian Network. Proceedings of the 2015 International Conference on Advanced Research in Computer Science Engineering & Technology (ICARCSET 2015), ACM.

[68] Lloyd, S. (1982). Least squares quantization in PCM. Information Theory, IEEE Transactions on 28(2): 129-137.

[69] Forgy, E. W. (1965). Cluster analysis of multivariate data: efficiency versus interpretability of classifications. Biometrics 21: 768-769.

[70] Patel, B. C. and D. G. Sinha (2010). An adaptive K-means clustering algorithm for breast image segmentation. International Journal of Computer Applications 10(4): 35-38.

[71] Utro, F. (2011). Algorithms for internal validation clustering measures in the Post Genomic Era. arXiv preprint arXiv:1102.2915.

[72] Redmond, S. J. and C. Heneghan (2007). A method for initialising the means clustering algorithm using trees. Pattern Recognition Letters 28(8): 965-973.

[73] Shirwaikar, R. and C. Bhandari (2013). K-means Clustering Method for the Analysis of Log Data. Int. Conf. on Advances in Signal Processing and Communication

[74] Gariel, M., A. N. Srivastava and E. Feron (2011). Trajectory clustering and an application to airspace monitoring. Intelligent Transportation Systems, IEEE Transactions on 12(4): 1511-1524.

[75] Vattani, A. (2011). K-means requires exponentially many iterations even in the plane. Discrete & Computational Geometry 45(4): 596-616.

[76] Akbari, E., Buntat, Z., Afroozeh, A., Pourmand, S. E., Farhang, Y., &Sanati, P. (2016). Silicene and graphene nano materials in gas sensing mechanism. RSC Advances, 6(85), 81647-81653.

[77] Ratanavilisagul, C. (2020, June). A novel modified particle swarm optimization algorithm with mutation for data clustering problem. In *2020 5th International Conference on Computational Intelligence and Applications (ICCIA)* (pp. 55-59). IEEE.

[78] Sinaga, K. P., & Yang, M. S. (2020). Unsupervised K-Means Clustering Algorithm. IEEE Access, 8, 80716-80727.

[79] Harifi, S., Khalilian, M., Mohammadzadeh, J., &Ebrahimnejad, S. (2020). Using Metaheuristic Algorithms to Improve k-Means Clustering: A Comparative Study. *Revue d'IntelligenceArtificielle*, *34*(3), 297-305.

[80] Xiaoqiong, W., & Zhang, Y. E. (2020). Image segmentation algorithm based on dynamic particle swarm optimization and K-means clustering. *International Journal of Computers and Applications*, *42*(7), 649-654.

[81] Janani, R., &Vijayarani, S. (2019). Text document clustering using spectral clustering algorithm with particle swarm optimization. *Expert Systems with Applications*, *134*, 192-200.

[82] C. Hua and W. Wei, "A Particle Swarm Optimization K-Means Algorithm for Mongolian Elements Clustering," 2019 IEEE Symposium Series on Computational Intelligence (SSCI), Xiamen, China, 2019, pp. 1559-1564, doi: 10.1109/SSCI44817.2019.9003077.

[83] Gupta, T., & Panda, S. P. (2019, February). Clustering Validation of CLARA and K-Means Using Silhouette & DUNN Measures on Iris Dataset. In *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)* (pp. 10-13). IEEE.

[84]A. Jaradat, B. Matalkeh and W. Diabat, "Solving Traveling Salesman Problem using Firefly algorithm and K-means Clustering," 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT), Amman, Jordan, 2019, pp. 586-589, doi: 10.1109/JEEIT.2019.8717463.

[85] S. Tiacharoen, "Adaptive K-means image segmentation based on meta heuristic algorithm," 2018 International Workshop on Advanced Image Technology (IWAIT), Chiang Mai, 2018, pp. 1-3, doi: 10.1109/IWAIT.2018.8369782.

[86] Das, P., Das, D. K., & Dey, S. (2017, December). PSO, BCO and K-means Based Hybridized Optimization Algorithms for Data Clustering. In 2017 International Conference on Information Technology (ICIT) (pp. 252-257). IEEE.