



## DISCOVERING OF OPTIMAL ELUCIDATION BY MEANS OF STOCHASTIC OPTIMIZATION

Zeyad Safaa Younus Alsaffawi<sup>1</sup>, Dzulkifli Mohamad<sup>2</sup>, Prof.Dr. Safaa Younus Alsaffawi<sup>3</sup>

<sup>1,2,3</sup> Faculty of Computing, Universiti Teknologi Malaysia, Faculty of Computer Sciences and Mathematics, University of Mosul, Mosul, Iraq

### ABSTRACT:

In several research communities, clustering explain methods for combination of unlabeled data and is used in assemblage, machine-learning circumstances, together with data mining, recovery of document and pattern organization. Efficiency of clustering algorithms depends mainly on the correctness of the resemblance determined to the data. The simple and the most well known clustering algorithms known is k-means algorithm and it remains as the most significant algorithm in the present days which commonly apply partitioned clustering algorithm. Being K-means algorithm the simple, quick and simple to combine with a variety of techniques, understandable and scalable it is mostly used in many applications for its enhanced performance in other larger systems. K means algorithm is flexible as well as simple however it has several limitations, the cluster consequence mostly depends on assortment of early cluster centroids and it might converge to local optima. A particle swarm is a population concerning particles, in which every particle is a moving entity which can move all the way through search space and is concerned to enhanced positions. Each particle consist of a position vector, which can stand for the candidate elucidation to problem, a velocity vector, as well as a memory vector, which is improved candidate solution encountered by a particle. Particle swarm optimization must contain a fitness assessment function to make a decision improved as well as best positions, function can obtain particle's position as well as assign it a suitability value.

**Keywords:** *K-means algorithm, Clustering, Particle swarm optimization, Elucidation, Data mining.*

## 1. INTRODUCTION:

Main purpose of clustering is to organize objects of data into various separate clusters basically as the intra cluster in which resemblance is maximum and other type in which the inter cluster difference among them is maximum. There have been numerous clustering algorithms available every year and the efficiency of algorithms depends on the aptness of the similarity measure to the data at hand [1][2]. Spatial data mining intend to mechanize knowledge detection procedure and plays a significant responsibility in extracting motivating spatial pattern in addition to characteristics; confine essential relationships connecting spatial as well as non-spatial information; present data constancy and at superior conceptual levels; and assist to rearrange spatial databases to attain improved performance. Cluster study has been applied ineffectively in past to common data mining in addition to machine learning. Feature assortment is procedure of identifying most effectual subset of innovative characteristics to make use of clustering. Feature extraction is employ transformations of input characteristics to construct novel salient description. Clustering as shown in fig1 is used in several research communities to

elucidate methods for the grouping of unlabeled data and is functional in numerous examining pattern-analyses, decision-making, and recovery of document, segmentation of image, and pattern organization [4][5]. Clustering can be used for allocating document storage in distributed systems. The simple and the most well known clustering algorithms known is k-means algorithm. It remains as the most significant algorithm in the present days which commonly apply partitioned clustering algorithm. It believes a Euclidean space and takes the quantity of clusters, k for granted. Particle swarm optimization has achieved a lot of consideration, and it has been functional in a lot of fields. Particle swarm optimization is a functional stochastic optimization algorithm on basis of population. The birds in a flock are corresponding to as particles, and particles are measured as simple agents flying all the way through a difficulty area [6]. K means algorithm is flexible as well as simple however it has several limitations, the cluster consequence mostly depends on assortment of early cluster centroids and it might converge to local optima. The similar initial cluster centre within a data space can constantly make similar cluster

consequence, if a good cluster centre can constantly be obtained; K-means will effort well. Advantages of K-mean clustering are simple as well as flexible; easy to understand along with implementation.

## 2. METHODOLOGY:

Cluster analysis is the association of a collection of patterns typically represented as a vector of dimensions, or a point in a multi dimensional space into clusters on the basis of resemblance. It is used for identification of similar groups of buyers, higher-level image processing: object detection, examination of traffic data in computer networks, revelation of complex graphs, text categorization in information recovery. The communities have different terminologies and statements for the components of the clustering practice and the contexts in which clustering are used. In several research communities, clustering explain methods for combination of unlabeled data and is used in assemblage, machine-learning circumstances, together with data mining, recovery of document and pattern organization. Hierarchical clustering algorithms construct a nested succession of partition on standard in support of integration or divide clusters basis on

resemblance. Partitional clustering system recognizes separation that optimizes a clustering standard. Latent semantic indexing is one of the efficient methods of spectral clustering which is aimed at discovering the most excellent subspace approximation to the original space of document by means of minimizing the Euclidean distance. Although methods of hierarchical clustering keep away from problem by systematizing document corpus to structure of hierarchical tree, clusters within every layer, on the other hand, do not unavoidably communicate to a significant combination of document corpus. In multi-dimensional problem space, particle's location can correspond to elucidation for difficulty. But Particle swarm optimization might lack global search capacity at end of a run due to exploitation of a linearly declining inertia weight as well as Particle swarm optimization might fail to discover the necessary optima when difficulty to be solved is moreover complex. Particle swarm optimization was based on swarming performance of animals as well as human social behaviour. Efficiency of clustering algorithms depends mainly on the correctness of the resemblance determined to the data. A particle swarm is a population

concerning particles, in which every particle is a moving entity which can move all the way through search space and is concerned to enhanced positions. Particle swarm optimization must contain a fitness assessment function to make a decision improved as well as best positions, function can obtain particle's position as well as assign it a suitability value. Each particle has individual coordinate as well as velocity to modify the flying direction in search space. And the entire particles move all the way through search space by following existing optimum particles. Each particle consist of a position vector, which can stand for the candidate elucidation to problem, a velocity vector, as well as a memory vector, which is improved candidate solution encountered by a particle. Particle swarm optimization based on intelligence and it is functional on both scientific research as well as engineering; Particle swarm optimization has no alteration as well as overlapping calculation; the search is taken place by speed of particle. The majority of optimist particle can capable to convey information onto previous particles during expansion of quite a lot of generations, and speed of researching is quicker. Particle swarm optimization accepts genuine number code,

and that is decided unswervingly by solution. Calculation in Particle swarm optimization is simpler as well as competent in global search. Being K-means algorithm the simple, quick and simple to combine with a variety of techniques, understandable and scalable it is mostly used in many applications for its enhanced performance in other larger systems. Components of a Clustering Task involves the following steps of pattern representation, definition of a pattern proximity determine appropriate to the data domain, clustering, data abstraction and evaluation of output.

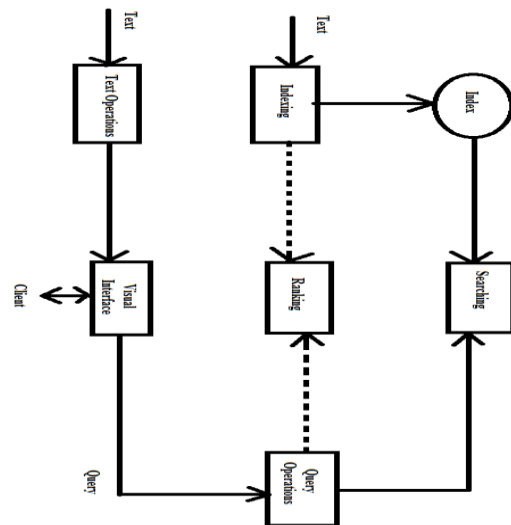


Fig1: An overview of clustering

### 3. CONCLUSION:

Clustering is used in several research communities to elucidate methods for the

grouping of unlabeled data and is functional in a numerous examining pattern-analysis, decision-making, and circumstances of machine-learning, together with data mining, assemblage, recovery of document, segmentation of image, and pattern organization. Hierarchical clustering algorithms construct a nested succession of partition on standard in support of integration or divide clusters basis on resemblance. A particle swarm is a population concerning particles, in which every particle is a moving entity which can move all the way through search space and is concerned to enhanced positions. Particle swarm optimization based on intelligence and it is functional on both scientific research as well as engineering; Particle swarm optimization has no alteration as well as overlapping calculation; the search is taken place by speed of particle. Being K-means algorithm the simple, quick and simple to combine with a variety of techniques, understandable and scalable it is mostly used in many applications for its enhanced performance in other larger systems.

## REFERENCES

- [1] Gabriela derban and Grigoreta sofia moldovan, "A comparison of clustering techniques in aspect mining", Studia University, Vol LI, Number1, 2006, pp 69-78.
- [2] H. M. Feng, C.Y. chen and F. Ye, "Evolutionay fuzzy particle swarm optimization vector quantization learning scheme in image compression", Expert Systems with Applications. Vol. 32, No. 1, 2007, pp. 213-222.
- [3] Jinxin D. And Minyong Q., "A new Algorithm for clustering based on particle swarm optimization and k-Means", International Conference Intelligence, 2009, pp 264-268.
- [4] Shalove Agarwal, Shashank Yadav and Kanchan Singh, "K-mean versus k-mean++ clustering Techniques", in IEEE 2012
- [5] Juntao Wang and Xiaolong Su, "An improved k-mean clustering algorithm", in IEEE, 2011, pp 44-46.
- [6] R. Eberhart and J. Kennedy, " Particle swarm optimization ", Proc. of the IEEE Int. Conf. on Neurad l Networks, Piscataway, NJ., 1995, pp. 1942-1948.