**CLUSTERING BASED OPTIMIZATION TECHNIQUES: A COMPARISON**

### A Project Work Synopsis

*Submitted in the partial fulfillment for the award of the degree of*

# BACHELOR OF ENGINEERING

### IN

### COMPUTER SCIENCE WITH SPECIALIZATION IN ARTIFICIAL INTELLIENCE AND MACHINE LEARNING

## Submitted by:

|  |  |
| --- | --- |
| **Ritika Chandel** | **Rajnandini Bhowmick** |
| **20BCS6556** | **20BCS6576** |
| **20AML4-A** | **20AML4-A** |

**Under the Supervision of:**

## Dr Anita



### CHANDIGARH UNIVERSITY, GHARUAN, MOHALI - 140413,

**PUNJAB**

#### MARCH, 2023

# 

# Abstract

# Optimization Wireless Sensor Network (WSN) is necessary to reduce redundancy and energy consumption. To optimizing wireless sensor networks for secured data transmission both at cluster head and base station data aggregation is needed. Data aggregation is performed in every router while forwarding data. The life time of sensor network reduces because of employing energy inefficient nodes for data aggregation. Hence aggregation process in WSN should be optimized in energy efficient manner. When sensors are deployed at different locations in wider area, it is possible to compromising attacks by adversaries. false data injected in compromised sensors during data aggregation process which results in false decision making at the Base Station (BS). Simple average data aggregation process is suitable only in attacker free environment. So to filter the false data during data aggregation, induced by the attacker. For every round of data aggregation need to observe the behavior of nodes. So that it easy to minimize an impact of attacker contribution at the final result. For secure data aggregation process along with trustworthiness estimation using Trust Weighted Secure Data Aggregation algorithm (TESDA). Data aggregation process is optimized by performing aggregation in energy efficient manner through clustering If the aggregator is compromised, then it affects entire aggregation accuracy. Hence it is necessary to propose a aggregation protocol that is resilient against compromised sensor and compromised aggregator in energy efficient and secure manner.

# *Keywords: optimization, redundancy, energy consumption, sensor network*

# Table of Contents

|  |  |  |
| --- | --- | --- |
|  | Title Page  Abstract | i |
| ii |
|  |
|  |
|  |
|
|
| **1.** | **INTRODUCTION** | **4** |
|  | * 1. Problem Definition   2. Research Overview/Specifications   3. Hardware Specification   4. Software Specification | 5  5  6  6 |
| 1. **LITERATURE SURVEY**    1. Existing Research    2. Proposed Research    3. Literature Review Summary 2. **PROBLEM FORMULATION** | | **6** |
| 6  10  10 |
| **13**  13  14  15  15  16 |
| 1. **RESEARCH OBJECTIVE** 2. **METHODOLOGY** 3. **EXPERIMENTAL SETUP** 4. **CONCLUSION** 5. **REFERENCES** | |

# INTRODUCTION

**Wireless Sensor Network (WSN)** is an infrastructure-less wireless network that is deployed in a large number of wireless sensors in an ad-hoc manner that is used to monitor the system, physical or environmental conditions.

Sensor nodes are used in WSN with the onboard processor that manages and monitors the environment in a particular area. They are connected to the Base Station which acts as a processing unit in the WSN System.   
Base Station in a WSN System is connected through the Internet to share data. Wireless sensor networks (WSNs) are a new class of wireless networks that are becoming very popular with a huge number of civilian and military applications. A wireless sensor network (WSN) is a wireless network that contains distributed independent sensor devices that are meant to monitor physical or environmental conditions. A WSN consists of a set of connected tiny sensor nodes, which communicate with each other and exchange information and data. These nodes obtain information on the environment such as temperature, pressure, humidity or pollutant, and send this information to a base station. The latter sends the info to a wired network or activates an alarm or an action, depending on the type and magnitude of data monitored [1–24].

Typical applications include weather and forest monitoring, battlefield surveillance, physical monitoring of environmental conditions such as pressure, temperature, vibration, pollutants, or tracing human and animal movement in forests and borders [1–23]. They use the same transmission medium (which is air) for wireless transmission as wireless local area networks (WLANs). For nodes in a local area network to communicate properly, standard access protocols like IEEE 802.11 are available. However, this and the other protocols cannot be directly applied to WSNs. The major difference is that, unlike devices participating in local area networks, sensors are equipped with a very small source of energy (usually a battery), which drains out very fast. Hence the need arises to design new protocols for MAC that are energy aware. Clearly there is some difference between a traditional WLAN and a WSN, as the latter has limited resources.

Sensor nodes offer a powerful combination of distributed sensing, computing and communication. The ever-increasing capabilities of these tiny sensor nodes, which include sensing, data processing, and communicating, enable the realization of WSNs based on the collaborative effort of a number of other sensor nodes. They enable a wide range of applications and, at the same time, offer numerous challenges due to their peculiarities, primarily the stringent energy constraints to which sensing nodes are typically subjected. As illustrated in Figure 1.1, WSNs incorporate knowledge and technologies from three different fields; Wireless communications, Networking and Systems and Control theory. In order to realize the existing and potential applications for WSNs, sophisticated and extremely efficient communication protocols are required. This chapter provides a first introduction to the WSNs, including architecture, specific characteristics and applications. The intrinsic properties of individual sensor nodes pose additional challenges to the communication protocols primarily in terms of energy consumption. As will be explained in the following chapters, WSN applications and communication protocols are mainly tailored to provide high energy efficiency. Sensor nodes carry limited power sources. Typically, they are powered through batteries, which must be either replaced or recharged (e.g., using solar power) when depleted. For some nodes, neither option is appropriate, that is, they will simply be discarded once their energy source is depleted. Whether the battery can be recharged or not significantly affects the strategy applied to energy consumption. Therefore, while traditional networks are designed to improve performance metrics such as throughput and delay, WSN protocols focus primarily on power conservation

.

#### Problem Definition

WSN has become one of the important technologies in the present decade. Energy consumption is the major challenge in the field of wireless sensor network. In WSN, there are some hard problems that cannot be solved in deterministic time. These hard problems can be solved by using optimization techniques. Clustering, routing, node localization, maintenance of the nodes, etc., are some of the hard problems that could be addressed. The main aim of these techniques is to provide the solution within specific time and also to minimize the consumption of the energy thus prolonging the lifetime of the network. This paper clearly describes the application of the different published optimization techniques in the field of WSN..

* 1. **Research Overview**

Saving energy and the communication of the data sensed by the nodes are two major issues in the wireless networks. Each node must consume little power and should work on low operating and system cost to maintain a large scale deployment of wireless sensor networks. Antenna and radio frequency transceiver are used for communication of sensor nodes with other nodes. Sensor nodes contain a memory unit, a CPU, the sensor unit, and the power source which is usually supplied by batteries. Many applications need sensor nodes to be designed as tiny as possible to create a small network suitable for any location. The small size of sensor nodes is beneficial for many situations but the small space also means the availability for battery capacity is small. A small network structure also provides another benefit of a reduced transmission range between nodes. The main focus for wireless sensor networks is mostly on energy conservation through various optimization techniques. These techniques should concentrate on communication and operation management as the power consumption needed for communication typically dominates a node’s power budget. In some applications there is a minimum need for sensing activity for most of the time and sometimes there is a need for strong sensor processing at particular instants. In such cases, there will be large variation of workloads. Energy awareness must be included into groups of communicating sensor nodes of the entire network as well as into the individual nodes in such applications. Wireless sensor networks have two important challenges in the deployment. These challenges include involvement of large number of devices and the need to embed them in a dynamic physical environment. The wireless sensor network capabilities involve mechanisms of distributing information over many nodes and collecting data to a sink node. These mechanisms should be low energy consuming, scalable with the number of nodes, and fault tolerant. Some sensor nodes may fail or paused due to lack of power, physical damage, or environmental obstruction. If there are multiple instances of failure occurring at a time, these mechanisms need to have the ability to form new links and routes to the sinks. This may include rerouting the packets or regulating the power to redirect the path with ample energy for processing.

* 1. **Hardware Specification**
* Processor: Intel Core i7
* RAM: 32GB
* GPU: NVIDIA GTX 1080 Ti with 11GB RAM
* Programming Language: Python
  1. **Software Specification**
* Jupyter Notebook
* VS code

1. **LITERATURE REVIEW**
   1. **Existing Research:**

**P. Lalwani, S. Das. “Bacterial Foraging Optimization Algorithm for CH selection and routing in wireless sensor networks.”**

The authors explore the application of a bio-inspired optimization algorithm known as the Bacterial Foraging Optimization Algorithm (BFOA) for solving the challenging problem of Cluster Head (CH) selection and routing in Wireless Sensor Networks (WSNs). The authors address the application of the Bacterial Foraging Optimization (BFO) algorithm to two critical aspects of Wireless Sensor Networks (WSNs): Cluster Head (CH) selection and routing. In WSNs, the selection of cluster heads and efficient routing are fundamental challenges that directly impact network performance and energy efficiency. The use of the BFO algorithm in this context is notable and signifies an attempt to leverage bio-inspired optimization techniques for addressing WSN issues

**A.A.Abba Ari, A. Gueroui, B. O. Yenke, N. Labraoui. “Energy efficient clustering algorithm for Wireless Sensor Networks using the ABC metaheuristic.**

The authors addressed the application of the Artificial Bee Colony (ABC) metaheuristic to the critical problem of energy-efficient clustering in Wireless Sensor Networks (WSNs). WSNs are characterized by resource-constrained sensor nodes, and energy efficiency is a paramount concern in their operation. The use of the ABC metaheuristic in this context signifies an attempt to leverage nature-inspired optimization techniques for addressing energy efficiency issues in WSNs.

**lhem Boussaïd, Julien Lepagnot, Patrick Siarry. “A survey on optimization metaheuristics.”**

The authors provide an in-depth overview of optimization metaheuristics. Metaheuristic algorithms are widely used for solving complex optimization problems in various domains,

and this survey paper aims to offer a comprehensive examination of these techniques. The authors give a comparative analysis of different metaheuristic algorithms in terms of their performance, convergence speed, and adaptability to various problem type. The research paper provides a comprehensive understanding of various metaheuristic algorithms, their applications, and their role in tackling challenging optimization problems.

**L. Aziz, S. Raghay, H. Aznaoui, A. Jamali. “A new approach based on a genetic algorithm and an agent cluster head to optimize energy in Wireless Sensor Networks.”**

The authors of this study presents a novel approach to address energy optimization in Wireless Sensor Networks (WSNs). The authors give a presentation of comparative results, potentially showing how the genetic algorithm and agent cluster head approach outperforms or compares to existing methods. Overall, the paper appears to introduce a novel approach that combines genetic algorithms and agent-based cluster heads to address energy optimization in WSNs. Such an approach has the potential to significantly improve the energy efficiency and overall performance of wireless sensor networks, making it a promising contribution to the field of WSN research

**“Steinley, D., & Brusco, M. J. (2007). Initializing k-means batch clustering: A critical evaluation of several techniques. *Journal of Classification, 24*(1), 99–121.**

The authors focus on the critical aspect of initializing the K-Means clustering algorithm. K-Means is a popular and widely used clustering technique that partitions data into K clusters based on similarity. However, the choice of initial cluster centroids can significantly impact the quality and efficiency of the clustering process. This paper provides a detailed examination of various techniques for initializing the K-Means algorithm and their effects on clustering results The research paper provides insights into the impact of initialization techniques on clustering quality and performance. Understanding the strengths and limitations of different initialization methods is crucial for achieving meaningful and reliable clustering results.

**“ Lloyd, S. P. (1957). Least squares quantization in PCM. Technical Report RR-5497, Bell Lab.**

This paper laid the foundation for a key concept in data encoding known as Lloyd-Max quantization, which is fundamental in various applications, including image and audio compression. The paper likely starts by introducing the concept of quantization in the context of Pulse Code Modulation (PCM), which is a method for digitally encoding analog signals. Quantization involves mapping a continuous range of input values (analog signal) to a finite set of discrete output values (quantized representation).Lloyd's work addresses the problem of finding the optimal quantization scheme that minimizes the mean squared error between the original analog signal and its quantized representation. This problem is crucial for achieving efficient data compression while maintaining acceptable signal quality. The paper discusses the practical implications and applications of Lloyd-Max quantization, especially in the context of digital signal processing and telecommunications

**“MacQueen, J. B. (1967). Some methods for classification and analysis of multivariate**

**observations. In L. M. Le Cam & J. Neyman (Eds.), *Proceedings of the fifth Berkeley***

***symposium on mathematical statistics and probability* (Vol. 1, pp. 281–297). California:**

**University of California Press**

The paper presented at the fifth Berkeley symposium on mathematical statistics and probability in 1967, is a seminal work in the field of data analysis and clustering. This paper introduced the K-Means clustering algorithm, a fundamental technique widely used for partitioning data into groups based on similarity. The paper likely starts by introducing the concept of clustering in the context of multivariate observations. Clustering involves grouping similar data points into clusters or groups while maximizing the dissimilarity between clusters. The paper may discuss how the K-Means algorithm compares to other clustering techniques available at the time. It likely highlights the algorithm's efficiency and its ability to handle large datasets.

**“Patrick J.F, Groenen, U. , Kayamak Fuzzy clustering with Minkowski distance functions”**

The report likely begins with an introduction to fuzzy clustering, a variant of traditional clustering where data points can belong to multiple clusters with varying degrees of membership. It may discuss the advantages of fuzzy clustering over hard clustering methods. The report likely addresses the problem of fuzzy clustering and introduces the specific problem or dataset that is being analyzed and clustered using Minkowski distance functions. Given that the Minkowski distance function has a parameter (p) that can be adjusted, the paper may discuss how the choice of this parameter affects the clustering results and how it can be tuned for different applications.

**“ Cai, J.Y., Xie, F.D., Zhang, Y.: Fuzzy c-means algorithm based on adaptive Mahalanobis distance. Comput. Eng. Appl. 174–176(2010)**

. The paper is likely to present experimental results and findings, demonstrating the effectiveness of the proposed approach in terms of clustering quality and accuracy. It may provide comparisons with other clustering methods or variations of fuzzy C-means.he core contribution of the paper is the development of an adaptive Mahalanobis distance metric. This adaptation likely takes into account the covariance structure of the data and adjusts the Mahalanobis distance during the clustering process. The adaptation might be based on the observed characteristics of the data.The report likely addresses the problem of fuzzy clustering and introduces the specific dataset or application context in which the authors propose to apply the fuzzy C-means algorithm with an adaptive Mahalanobis distance.

**“Comparison of various landmark detecting techniques in the context of forensic facial recognition” (Pool, 2018)**

This study compares the performance of Dlib and STASM, two distinct landmark identification algorithms. First, quantitative research was conducted using landmarks. For Second, a qualitative analysis was conducted in which the placement of the landmarks placed by human examiners and those discovered by Dlib and STASM were contrasted (Pool, 2018). In 7,6% of all photos, STASM identified a face where there was none. As a result, Dlib outperformed STASM in nearly every case in the quantitative analysis (Pool, 2018). So, the STASM landmarks are less accurate compared to Dlib landmarks according to the human examiners. (Pool, 2018)

**“Facial feature point detection: A comprehensive survey” (Wang et al., 2018)**

Utilizing the help of a large number of manually annotated photos, this work provides a thorough assessment of facial feature point recognition. Numerous applications benefit from facial feature point detection (Wang et al., 2018). There are two primary kinds of current approaches: model based on parametric shape methods, and model based on nonparametric shape methods, depending on whether a parametric form model is required. The parametric techniques are further divided into two secondary types depending on their appearance models: models based on local component methods and models based on comprehensive methods.

**Proposed System:**

In this research paper, we aim to study and compare the early and modern clustering techniques that is used as an optimization techniques in wireless sensor network

**.**

* 1. **Literature Review Summary:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Year** | **Article/Author** | **Technique** | **Source** | **Evaluation Parameter** |
| 2022 | P. Lalwani, S. Das. “Bacterial Foraging Optimization Algorithm for CH selection and routing in wireless sensor networks.” 2016 3rd International Conference on Recent Advances in Information Technology (RAIT), Dhanbad, 2016: pp. 95–100. | Using the optimization parameters for wireless sensor network | *Journal of Computer Science & Technology*, vol. 22, no. 1, pp. 12–41, 2022 | Clustering-based-optimization technique in wireless sensor network |
| 2022 | G. Gajalakshmi, G. U. Srikanth. “A survey on the utilization of Ant Colony Optimization (ACO) algorithm in WSN. | Utilization of ACO optimization algorithm in WSN | *IEEE Transactions on Affective Computing* (2022) 13(3) 1195-1215 | Shortes Patht echniques using ACO |
| 2020 | A.A.Abba Ari, A. Gueroui, B. O. Yenke, N. Labraoui. “Energy efficient clustering algorithm for Wireless Sensor Networks using the ABC metaheuristic.” | Energy efficient clustering algorithm for wireless sensor network | *Electronics (Switzerland)*. | Efficient clustering algorithms |
|  |  |  | MDPI AG, pp. 1–53. |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| 2019 | L. Aziz, S. Raghay, H. Aznaoui, A. Jamali. “A new approach based on a genetic algorithm and an agent cluster head to optimize energy in Wireless Sensor Networks | Efficient use of genetic algorithm in wireless sensor ne | *International Journal of Computer Network*, 127(2),  pp. 115–142. | Node energy consumptions( Minimization) |
| 2019 | P. Lalwani, I. Ganguli, H. Banka. “FARW: Firefly algorithm for Routing in wireless sensor networks.” | Firefly algorithm for routing for wireless sensor network | *Big Data and Cognitive computing* | Comparison with conventional algorithms like EADC, DHCR and hybrid routing protoicol |
|  |  |  | . |  |
|  |  |  | . |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| 2016 | A. Singh, S. Rathkanthiwar, S. Kakde. “Energy efficient routing of WSN using particle swarm optimization and V-LEACH protocol.” | Efficient use of routing techniques in wireless sensor network. | *International Journal of Advanced Trends in Computer Science and Engineering*, 8(1.6 Special  Issue), pp. 525–530. | Performance comparison with existing leach protocol shows proposed protocol provides better performance to minimize energy dissipation in the transmission |
| 2019 | Rakesh Kumar Yadav, Akshita Kapila, Deepa Arora. “Analysis Geographic Routing Protocol using Ant Colony Optimization Technique for WSNs. | Compared the effective optimization algorithms | *Applied Sciences (Switzerland)*. MDPI AG. | Accuracy |
| 2018 | Sandıkcı, E.N., Erdem, Ç.E. and Ulukaya, S. (2018) “Optimization techniques Methods” (Sandıkcı,  Erdem, and Ulukaya, 2018) | Analysis of the most effective optimization techniques. | *26th Signal Processing and Communications Applications Conference (SIU),* 2018. | Energy consumption |
| 2018 | Bingnan Pei, Hao Zhang, Tengda Pei, H. Wang. “Firefly algorithm optimization based WSN localization algorithm.” | Optimization technique using firefly algorithm | *Journal of University of Twente,* 2018. | Modification of the DV-Hop optimization algorithm(by correcting the position of the unknown nodes) |

# PROBLEM FORMULATION

Wireless sensor Networks or WSN have a varying range of applications in deployment from military areas where it is used extensively for remote data collection, triangulating coordinates, surveillance, to biological and agriculture infrastructure in crop monitoring and estimation of geophysical parameters. WSN can utilize multi-hop communication technique by communication with each other directly or through relay nodes. Remote sensing and wireless sensors networks have become widely used in many critical operations like reconnaissance on inaccessible altitudes and hostile terrains. Depending upon the range and feasibility, they can use their properties to benefit from clustering and optimization algorithms. The design of WSNs however faces many challenges. WSN’s are restricted to limited bandwidth and energy as well as low processing power to include with. As such the algorithms for WSNs must optimize all of the design constraints to obtain the best possible results. WSNs are data accumulators. It gathers loads of data from various nodes while avoiding redundancy and also maintains the integrity of the data. In this work, different novel optimization techniques have been explored with regard to WSNs Wireless sensor networks (WSNs) is the collection of various types of sensor nodes that performs various tasks in a collaborative way. It is used healthcare, military, disaster management system, etc. In the research article various optimization techniques such as particle swarm optimization, gravitational search algorithm, harmony search algorithm and many others. and NP- Complete problems addressed by various authors in WSNs were presented mainly relay node placement and CH selection. Major objective of this research article was to address various findings in WSNs and how to solve these issues using optimization techniques. In addition, this research article was also provide future direction of research in WSN

# OBJECTIVES :

In this research paper, we aim to study and compare the early and modern clustering techniques that is used as an optimization techniques in wireless sensor network. Clustering is one of the important methods for extend the network lifetime in wireless sensor networks (WSNs). It involves grouping of sensor nodes into clusters and electing cluster heads (CHs) for all the clusters. Wireless sensor networks (WSN) are spatially distributed separate sensors to monitor physical or environmental conditions, like temperature, sound, pressure, etc. as well as to cooperatively push through their data through the network to a base station. The WSN is built of nodes from a few to several hundreds or even thousands, where each node is connected to each other sensors. Clustering is one of the important methods for extend the network lifetime in wireless sensor networks (WSNs). It involves grouping of sensor nodes into clusters and electing cluster heads (CHs) for all the clusters. CHs collect the data from particular clusters nodes and forward the aggregated data to base station, so in order to energy efficient clustering is well known optimization problem which has been calculated widely to extend lifetime of wireless sensor networks (WSNs) and there are many types of clusters

Fuzzy C-Means Clustering. Fuzzy c-means (FCM) is a data clustering technique where each data point belongs to a cluster to a degree that is specified by a membership grade. The FCM algorithm starts with an initial guess for the cluster centers, which represent the mean location of each cluster. **Fuzzy Clustering** is a type of clustering algorithm in machine learning that allows a data point to belong to more than one cluster with different degrees of membership. Unlike traditional clustering algorithms, such as k-means or hierarchical clustering, which assign each data point to a single cluster, fuzzy clustering assigns a membership degree between 0 and 1 for each data point for each cluster

In **K-means Clustering** We are given a data set of items, with certain features, and values for these features (like a vector). The task is to categorize those items into groups. To achieve this, we will use the K-means algorithm; an unsupervised learning algorithm. ‘K’ in the name of the algorithm represents the number of groups/clusters we want to classify our items into. The algorithm will categorize the items into k groups or clusters of similarity. To calculate that similarity, we will use the Euclidean distance as a measurement.

**METHODOLOGY**

For this paper, we seek to test the early and modern clustering techniques such as the K-means clustering and the fuzzy- C means clustering to conduct a comparative study about the strengths and weakness of each of the methods and to find out which method outperforms others under which parameters. We seek to complete this study following the following steps:

1. Create in new Virtual Environment in Jupyter notebook
2. Downloading all the necessary libraries in the new virtual environment such as – Python, SciPy, NumPy, OpenCV, matplotib. Sklearn, KMeans etc.
3. Then creating an environment by imporing all the necessary libraries and testing the methods
4. Importing the library of the method to be tested.
5. Finding out the the efficiency, accuracy and speed of each algorithm
6. Comparing the results and finding the conclusion.

**EXPERIMENTAL SETUP**

To conduct this study, the required setup is very minimal. One needs a PC with Jupyter Labs, VS Code , the libraries containing the methods to be compared and then perform the implementation to find out the required result

**CONCLUSION**

The performance of the fuzzy c-means algorithm gives better performance than k-mean, both when using thresholding with mean and median methods. Better performance of fuzzy c-means requires additional time when compared to k-means. the computational time of fuzzy c-means is longer than k-means. This is also supported by the complexity of the k-mean Q(n) algorithm, while fuzzy c-means Q(n2).[K-Means clustering](http://en.wikipedia.org/wiki/K-means_algorithm) and [Fuzzy-C Means Clustering](http://en.wikipedia.org/wiki/Data_clustering#Fuzzy_c-means_clustering) are very similar in approaches. The main difference is that, in Fuzzy-C Means clustering, each point has a weighting associated with a particular cluster, so a point doesn't sit "in a cluster" as much as has a weak or strong association to the cluster, which is determined by the inverse distance to the center of the cluster. Fuzzy-C means will tend to run slower than K means, since it's actually doing more work. Each point is evaluated with each cluster, and more operations are involved in each evaluation. K-Means just needs to do a distance calculation, whereas fuzzy c means needs to do a full inverse-distance weighting.

In this project, we detected the presence of Parkinson’s

# REFERENCES

1. [k means clustering required libraries - Google Search](https://www.google.co.in/search?q=k+means+clustering+required+libraries&sca_esv=573962864&sxsrf=AM9HkKm3qaDgeP_9Llvfr11-73QsC8jspw%3A1697516935857&ei=hw0uZeTfM-fz4-EPttulsA0&oq=k+means+clustering+required+libra&gs_lp=Egxnd3Mtd2l6LXNlcnAiIWsgbWVhbnMgY2x1c3RlcmluZyByZXF1aXJlZCBsaWJyYSoCCAAyBRAhGKABMggQIRgWGB4YHUiIuwVQwQVYibAFcAF4AZABAJgB4wKgAcUqqgEIMC4yOC40LjG4AQHIAQD4AQGoAgrCAgcQIxjqAhgnwgIEECMYJ8ICBxAjGIoFGCfCAggQABiKBRiRAsICBxAuGIoFGEPCAgcQABiKBRhDwgINEC4YigUYxwEY0QMYQ8ICDhAAGIoFGLEDGIMBGJECwgIFEAAYgATCAg0QABiKBRixAxiDARhDwgILEC4YigUYsQMYgwHCAgsQLhiABBixAxiDAcICEBAAGIAEGBQYhwIYsQMYgwHCAgsQABiABBixAxiDAcICCxAAGIoFGLEDGJECwgIIEAAYigUYhgPCAgYQABgWGB7CAgcQIRigARgKwgIEECEYFcICChAhGBYYHhgPGB3iAwQYACBBiAYB&sclient=gws-wiz-serp)
2. [Introduction to k-Means Clustering with scikit-learn in Python | DataCamp](https://www.datacamp.com/tutorial/k-means-clustering-python).
3. [Introduction to k-Means Clustering with scikit-learn in Python | DataCamp](https://www.datacamp.com/tutorial/k-means-clustering-python)
4. <https://doi.org/10.3390/app9214678>.
5. [K means Clustering - Introduction - GeeksforGeeks](https://www.geeksforgeeks.org/k-means-clustering-introduction/)
6. https://stackoverflow.com/questions/2345903/whats-is-the-difference-between-k-means-
7. [K-Means Clustering in Python: A Practical Guide – Real Python](https://realpython.com/k-means-clustering-python/).
8. [Fuzzy Clustering - MATLAB & Simulink (mathworks.com)](https://www.mathworks.com/help/fuzzy/fuzzy-clustering.html)
9. [Fuzzy clustering - Wikipedia](https://en.wikipedia.org/wiki/Fuzzy_clustering)
10. [Unit-2.pdf (egyankosh.ac.in)](https://egyankosh.ac.in/bitstream/123456789/12678/1/Unit-2.pdf)
11. [C-Means Clustering Explained | Built In](https://builtin.com/data-science/c-means)
12. [Fuzzy Clustering - an overview | ScienceDirect Topics](https://www.sciencedirect.com/topics/computer-science/fuzzy-clustering)
13. [Data Clustering Algorithms - Fuzzy c-means clustering algorithm (google.com)](https://sites.google.com/site/dataclusteringalgorithms/fuzzy-c-means-clustering-algorithm)