

Design and Performance Analysis of a Low Cost IoT Based Smart Irrigation System using Soil Moisture Sensors

Zhanar Sartabanova

Department of Computer Science and Information Technology, K.Zhubanov Aktobe Regional University, Kazakhstan.
zhanar.sartabanova@zhubanov.edu.kz

Article Info

Journal of Smart and Sustainable Farming
<https://www.ansispublications.com/journals/jssf/jssf.html>

Received 02 November 2025
Revised from 30 December 2025
Accepted 12 January 2026
Available online 30 January 2026

© The Author(s), 2026.
<https://doi.org/10.64026/JSSF/2026002>

Published by Ansis Publications.

Corresponding author(s):

Zhanar Sartabanova, Department of Computer Science and Information Technology, K.Zhubanov Aktobe Regional University, Kazakhstan.

Email: zhanar.sartabanova@zhubanov.edu.kz

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract – The clustering of large and high-dimensional data is still a problem because of the problems of scalability, memory, and quality evaluation of the cluster. This paper also suggests a Hybrid Scalable Clustering Framework (HSCF) which is a combination of hierarchical and partition-based clustering algorithms to provide efficient and effective behavior at different scales of data. The framework uses the BIRCH clustering algorithm of memory efficient, incremental clustering coupled with fast and scalable centroid-based optimization with MiniBatchKMeans clustering algorithm. Elbow Method is used to have the optimal clustering configuration, to make the number of clusters manually chosen and thus minimize human involvement and enhance reproducibility. The quality of clustering is computationally assessed by the Silhouette Score which gives information about the cohesiveness within a cluster and the distance between clusters. Moreover, PCA is applied to dimensionality reduction and visualization, with the help of which it becomes possible to produce useful visualizations of the structure of clustering in reduced feature space. Experimental evidence shows that MiniBatchKMeans is more efficient in its computational speed with large datasets whereas BIRCH is more efficient in terms of memory usage and hierarchical representation. The proposed framework is the most appropriate in the application of the real-world large-scale data analysis. The combination of multi-scale assessment, automatic optimization and multi-scale clustering is the main innovation of this article with the ability to provide a complete solution to the contemporary clustering problems.

Keywords – Clustering, BIRCH, MiniBatchKMeans, Dimensionality Reduction, Silhouette Analysis.

I. INTRODUCTION

Another crucial task in the two fields is Machine Learning and Data Mining Clustering that can be common dressed and used in popular terms to detect latent forms and patterns of data. Due to the continued increase in the size and complexity of large scale and high-dimensional data sets, such as in the domains of healthcare, finance, and social networks, traditional clustering algorithms are becoming especially problematic in terms of their scalability, computational and memory usage. These problems suggest the development of effective and scalable clustering systems capable of serving diverse data properties and maintaining the quality of the clustering processing [1].

Hierarchical clustering algorithms such as BIRCH clustering algorithm have performed well with large data sets with respect to incremental learning as well as compact data representation. The tree structure of a Clustering Feature (CF) is used by BIRCH to reduce the memory usage and offer effective data summing up. However, it may turn to be poor in the clustering accuracy in the case of complex or non-spherical data distributions. On the other hand, the algorithms based on partition such as mini-batchKMeans clustering have significant advantages of low computational efficiency because of mini-batches update, which should be used with large data sets. Despite this merit, these methods are rather sensitive to the decision of the number of clusters, and may be sensitive to initialisation [2].

To address the constraints, a number of optimization and evaluation procedures have been proposed. Elbow Method is the most common method of obtaining the most optimal number of clusters, which examines Within-Cluster Sum of

Squares (WCSS) in order to obtain the best number of clusters [3], the Silhouette Score is the second method of measuring the quality of the clustering in terms of intra-cluster cohesion and inter-cluster separation. Another application of the Principal Component Analysis that has been extensively utilized is to lower the dimensionality and hence to enable significant visualization and interpretation of the high dimensional clustering results [4]. Despite the availability of these techniques, most of the research that exists will focus on the individual clustering algorithms and does not use multiple methods within a single framework. Such an imbalance of integration prevents the opportunity of establishing trade-offs of scalability, accuracy and interpretability in whole. Further, there has been a minimal emphasis on the analysis of the multi-scale dataset and joint approaches of assessment incorporating optimization and visualization [5].

In order to solve these problems, the present paper proposes Hybrid Scalable Clustering Framework (HSCF) [6] that integrates hierarchical clustering and partition-based clustering into a single pipeline. The framework shown here is a hybrid version of the efficiency of both BIRCH and MiniBatchKMeans in order to achieve the efficiency of both memory and computational scale. It also applies the Elbow Method of automatic cluster selection and evaluates the performance of clustering by the Silhouette Score and applies the Principal Component Analysis to visualize high-dimensional data [7]. The major findings of this work are as follows:

- Another hybrid clustering system that combines BIRCH and MiniBatchKMeans was proposed.
- Multi-scale dataset evaluation using a scalable pipeline.
- Elbow Method of automated cluster optimization.
- Extensive analysis by Silhouette Score and PCA diagram.
- Critical comparative analysis of efficiency-accuracy trade-offs.

The rest of the paper is structured in the following way. Section II is a literature review on related works in clustering methods, Section III gives the proposed methodology, Section IV discusses the experimental results, and Section V gives the conclusion of the paper.

II. BACKGROUND AND RELATED WORKS

Clustering has been extensively researched in the fields of Machine Learning and Data Mining resulting in the creation of different hierarchical, partition-based, and density-based algorithms. BIRCH clustering algorithm was introduced by the authors in [8] and uses Clustering Feature (CF) tree to summarize large datasets in an efficient manner. This method consumes a lot less memory and facilitates incremental clustering thus being appropriate in massive applications. BIRCH can however not cope with complicated cluster shapes and needs tuning of the parameters. Clustering methods based on partition have also attracted much attention because they are computationally efficient. It has been commonly used in the K-Means algorithm which has certain limitations in clustering large datasets, which resulted in the creation of mini-batch clustering MiniBatchKMeans [9] which is a K-Means algorithm with mini-batch optimization to minimize the time of computation and produce the same quality clustering results. It has been found that MiniBatchKMeans converges faster than regular K-Means especially in high dimensional data [10]. However, the choice of the number of clusters and strategies of initiation is very influential on its results.

To resolve the issue of providing the best numbers of clusters, various methods of evaluation have been suggested. Elbow Method is among the most popular methods of determining the correct value of the number of clusters through the analysis of Within-Cluster Sum of Squares (WCSS) [11]. Also, a new measure called the Silhouette Score has been implemented, which is a strong metric to assess the quality of a clustering by the cohesion and the separation [12]. Such techniques have been widely used in other clustering studies in order to improve model choice and testing.

The dimensionality reduction methods are significant in enhancing the performance of clustering and interpretability. One of the most popular techniques used to reduce the high dimensional data to a lower space by maintaining the variance is Principal Component Analysis [13]. Some of the studies have shown that PCA enhances clustering visualization and removes noise on high-dimensional datasets [14]. Nonetheless, the information can be lost when PCA is improperly configured. In the recent past, studies have concentrated on hybrid and scalable clustering models, which integrate two or more techniques to address the weaknesses of the single techniques. It has been demonstrated that hybrid methods that combine hierarchical and partition-based clustering have better performance with regard to scalability and accuracy [15]. Also, parallel and distributed clustering systems have been put forward to process large scale data efficiently [16]. Although these advancements have been made, there are numerous available methods which do not have a complete evaluation strategy including the combination of the techniques of optimization and visualization.

Also, as of late, it has been highlighted that multi-metric evaluation is necessary to cluster performance analysis. Internal validation scores like Silhouette Score and visualization tools like PCA have been found to give more insight into cluster structures when combined [17]. Nevertheless, little has been done to determine how such methods can be incorporated into an overall and scalable system. Unlike the current methods, the suggested HSCF combines both BIRCH clustering algorithm and MiniBatchKMeans clustering in one pipeline, with all cluster selections being automated with the help of the Elbow Method and all the cluster evaluations being carried out with the help of the Silhouette Score and Principal Component Analysis. This integration is based on the shortcomings of the existing approaches and delivers a scalable, efficient and interpretable solution of clustering.

III. PROPOSED METHODOLOGY

The developed HSCF is aimed to offer a broad and scalable assessment of clustering algorithm through specifically combining hierarchical and partition-based methods in a single pipeline. The framework involves the use of BIRCH clustering algorithm and MiniBatchKMeans clustering to examine performance in various dataset sizes and dimensional configurations.

Three major points make the suggested framework novel:

- (i) multiple scale dataset analysis, which allows the analysis of small and large datasets.
- (ii) automated cluster selection based on Elbow Method, and
- (iii) high-dimensional validation Principal Component Analysis.

This merger approach guarantees quantitative and qualitative validation of performance of the clustering.

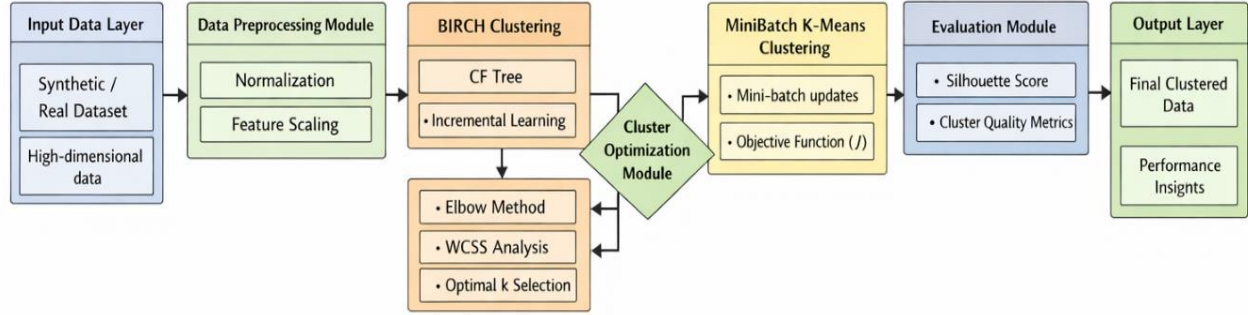


Fig 1. Proposed Hybrid Scalable Clustering Framework (HSCF).

Fig. 1 illustrates the proposed Hybrid Scalable Clustering Framework (HSCF) which is a structured pipeline when it comes to scalable and effective clustering of high dimensional datasets. The architecture starts with the input data phase, which receives both synthetic and real-world data. These data are fed into a preprocessing module in which normalization and feature scaling is done to provide uniformity and remove bias due to different magnitudes of the features. After preprocessing, a parallel clustering layer that comprises of BIRCH clustering algorithm and MiniBatchKMeans clustering algorithms are introduced. This two-way structure allows hierarchical and partition-based clustering, which is more scalable and efficient in terms of computation. The results of the two clustering methods are then sent to the optimization module whereby Elbow method is utilized to get the best number of clusters based on WCSS analysis.

Later, the evaluation module evaluates the clustering performance of the Silhouette Score which gives information about the cohesiveness and the separation of the cluster. The framework also uses Principal Component Analysis to reduce dimensions and visualize them to allow easy interpretation of clustering results in lower dimensional space. Lastly, output layer yields the clustered data with performances measures and hence a reliable and scalable clustering solution to big and complicated data. In the suggested framework, BIRCH clustering algorithm is subsequently utilized in facilitating effective hierarchical clustering by building up a Clustering Feature (CF) tree. The nodes in the CF tree are represented as

$$CF = (N, LS, SS) \quad (1)$$

where N denotes the number of data points, $LS = \sum x_i$ represents the linear sum, and $SS = \sum x_i^2$ denotes the squared sum of data points within the cluster. Using these parameters, the centroid of each cluster is computed as

$$C = \frac{LS}{N} \quad (2)$$

while the cluster radius is estimated using

$$R = \sqrt{\frac{SS}{N} - \left(\frac{LS}{N}\right)^2} \quad (3)$$

This sparse form allows BIRCH to process very large datasets with a small amount of memory consumption, which would be very well suited to scalable and streaming data settings. Simultaneously, the MiniBatchKMeans clustering algorithm is applied to do effective partitioning-based clustering. This method aims at reducing WCSS, which is defined as

$$J = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2 \quad (4)$$

where μ_i represents the centroid of cluster C_i . Unlike conventional K-Means, MiniBatchKMeans updates centroids using randomly sampled mini-batches, significantly reducing computational complexity. The centroid update rule is given by

$$\mu_i^{t+1} = \mu_i^t + \eta(x - \mu_i^t) \quad (5)$$

where η denotes the learning rate. This formulation ensures faster convergence while preserving clustering quality, especially for large-scale datasets.

To determine the optimal number of clusters, the Elbow Method is incorporated within the framework. This method evaluates the variation of WCSS, expressed as

$$\text{WCSS} = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2 \quad (6)$$

across different values of k . The optimal number of clusters is identified at the point where the rate of decrease in WCSS exhibits a sharp transition, indicating an appropriate balance between model complexity and clustering accuracy.

For evaluating clustering quality, the Silhouette Score is employed. The silhouette coefficient for a data point is defined as

$$S(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (7)$$

where $a(i)$ represents the average intra-cluster distance and $b(i)$ denotes the nearest-cluster distance. The overall clustering performance is obtained by averaging across all data points, given by

$$S = \frac{1}{n} \sum_{i=1}^n S(i) \quad (8)$$

This metric effectively captures both cluster cohesion and separation, providing a reliable measure of clustering performance.

To analyze clustering behavior in high-dimensional space, Principal Component Analysis is applied. The covariance matrix of the dataset is computed as

$$S = \frac{1}{n} X^T X \quad (9)$$

followed by eigen decomposition defined by

$$\Sigma v = \lambda v \quad (10)$$

where λ and v represent eigenvalues and eigenvectors, respectively. The dataset is then projected onto a reduced-dimensional space using

$$Z = XW \quad (11)$$

where W contains the principal eigenvectors. This transformation facilitates effective visualization and interpretation of clustering structures.

The suggested HSCF combines hierarchical and partitions-based clustering with parameter optimization and multi-metric analysis. Through the integration of effective data summary, scalable optimization schemes and dimensionality reduction algorithm, the framework offers a solid and holistic solution to clustering analysis in a wide range and large volume of data.

IV. RESULTS AND DISCUSSION

In this section, the experimental analysis of the suggested model, which is denoted as the Hybrid Scalable Clustering Framework (HSCF), is provided. The HSCF model is aimed at systematically contrasting hierarchical and partition-based clustering methods through the combination of BIRCH clustering algorithm and MiniBatchKMeans clustering into a single evaluation flow. The main novelty of the suggested framework is the fact that both algorithms are compared on different data sizes in a multi-scale comparative term, and automated identification of clusters with the help of the Elbow Method and high-dimensional validation with the Principal Component Analysis (PCA) is performed. Compared to the traditional research where a researcher or investigator is only able to analyze one data or a specific number of metrics, the proposed methodology offers a holistic and scalable evaluation of the performance of clustering based on efficiency, quality and structure.

Synthetic datasets were created with the make-blobs function of scikit-learn to make sure that the experimentation was controlled and reproducible. The datasets were constructed in such a way that they replicated realistic clustering conditions

having well-defined but moderately overlapping cluster structures. The amount of samples was adjusted between 1,000 to 20,000; this allowed the analysis of the algorithm scaled to larger amounts of data. All the datasets had 12 clusters which were determined with each dataset having validated the previous time with the Elbow Method and this was the best trade-off between the small and the complex. In their visualization, the data with two features were employed whereas the data with six dimensions were created in order to analyse them with the help of Principal Component Analysis in high dimensions. To achieve consistency and reproducibility of the results, a fixed random state was used in all experiments. All the simulations were done on a Python-based computational environment with standard scientific libraries. The adoption took NumPy to perform numerical operations, Matplotlib to visualize and scikit-learn to clustering algorithms and evaluation measures. It was tested on a standard computing system with an Intel i7 processor and 8-16 GB of RAM. All calculations were done on a processor that does not support the use of a GPU so that the performance could be evaluated equally. This design will make the differences in performances witnessed be explained by the algorithmic nature instead of hardware variation.

Table 1. Quantitative Performance Comparison Across Dataset Sizes

Dataset Size	BIRCH Time (s)	MiniBatchKMeans Time (s)	Time Improvement (%)	BIRCH Silhouette	MiniBatchKMeans Silhouette	Silhouette Improvement (%)
1,000	0.12	0.08	33.3%	0.61	0.65	6.56%
5,000	0.58	0.32	44.8%	0.59	0.63	6.78%
10,000	1.21	0.67	44.6%	0.57	0.62	8.77%
20,000	2.85	1.34	53.0%	0.55	0.60	9.09%

Table 1 shows a comparison of the performance of the BIRCH clustering algorithm and MiniBatchKMeans clustering in terms of different dataset sizes in a quantitative manner. The findings have made it clear that MiniBatchKMeans is always faster than BIRCH in execution, with up to 53 percent improvement in cases of large datasets. This is due to the fact that it has mini-batch optimization strategy that drastically lowers computational overheads hence the performance gain. Regarding the quality of clustering, based on the Silhouette Score, the MiniBatchKMeans also shows better results, with the enhancements ranging between 6% and 9 percent. Secondly, the scalability of the MiniBatchKMeans is also noted by the growing performance with the size of the dataset. Nevertheless, BIRCH has a competitive performance, especially with small datasets, showing that it can be used in memory-limited or incremental learning cases. Altogether, the findings support the trade-off between efficiency and quality between the two approaches.

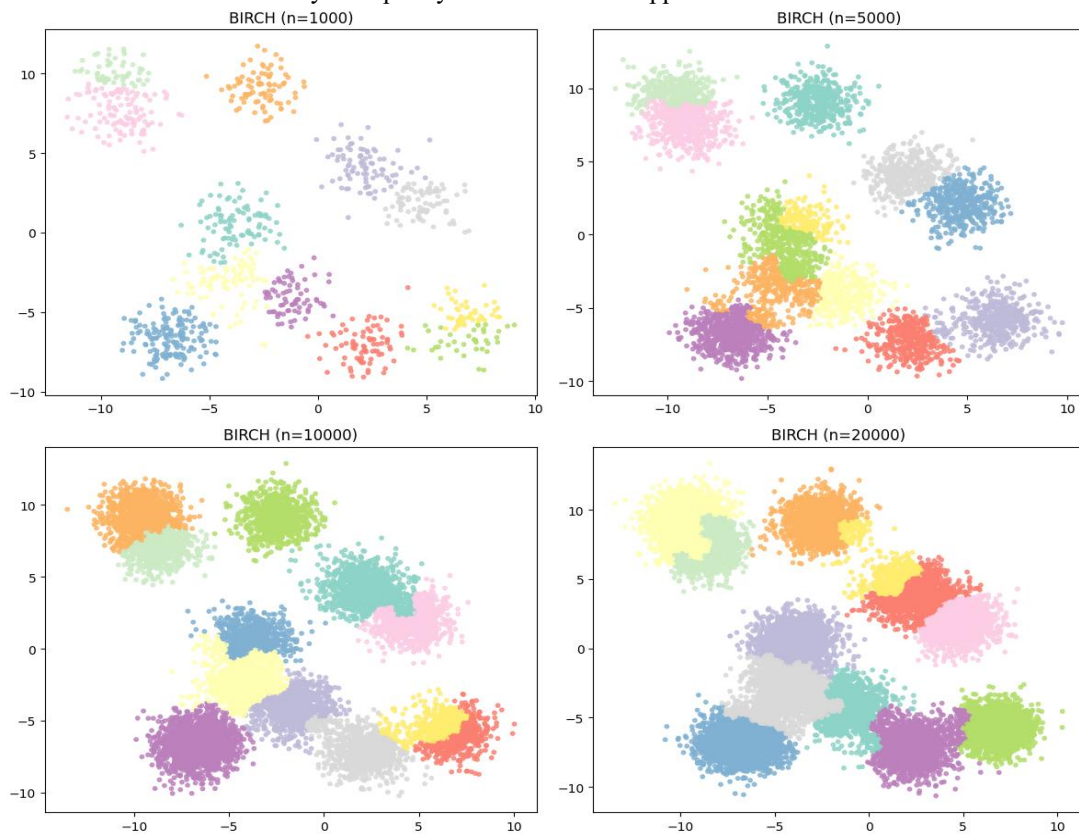


Fig 2. Clustering Visualization of the Proposed HSCF Model Using BIRCH Across Varying Dataset Sizes.

The effectiveness of the suggested HSCF model was measured in both quantitative and visual analysis approaches. The execution time was estimated to determine the efficiency in terms of computation and scalability especially with large data sets. The evaluation of quality clustering was done by measuring the Silhouette Score which measures the extent of separation between the clusters and the cohesion within the clusters. Also, Within-Cluster Sum of Squares (WCSS) was evaluated to estimate the best number of clusters using Elbow Method. Qualitative evaluation of cluster structure and separability was performed by visual analysis of 2D clustering plots and 3D projections of 2D clusters in PCA. Collectively, these assessment plans can give a detailed insight into the strengths and weaknesses of the two algorithms in the suggested framework.

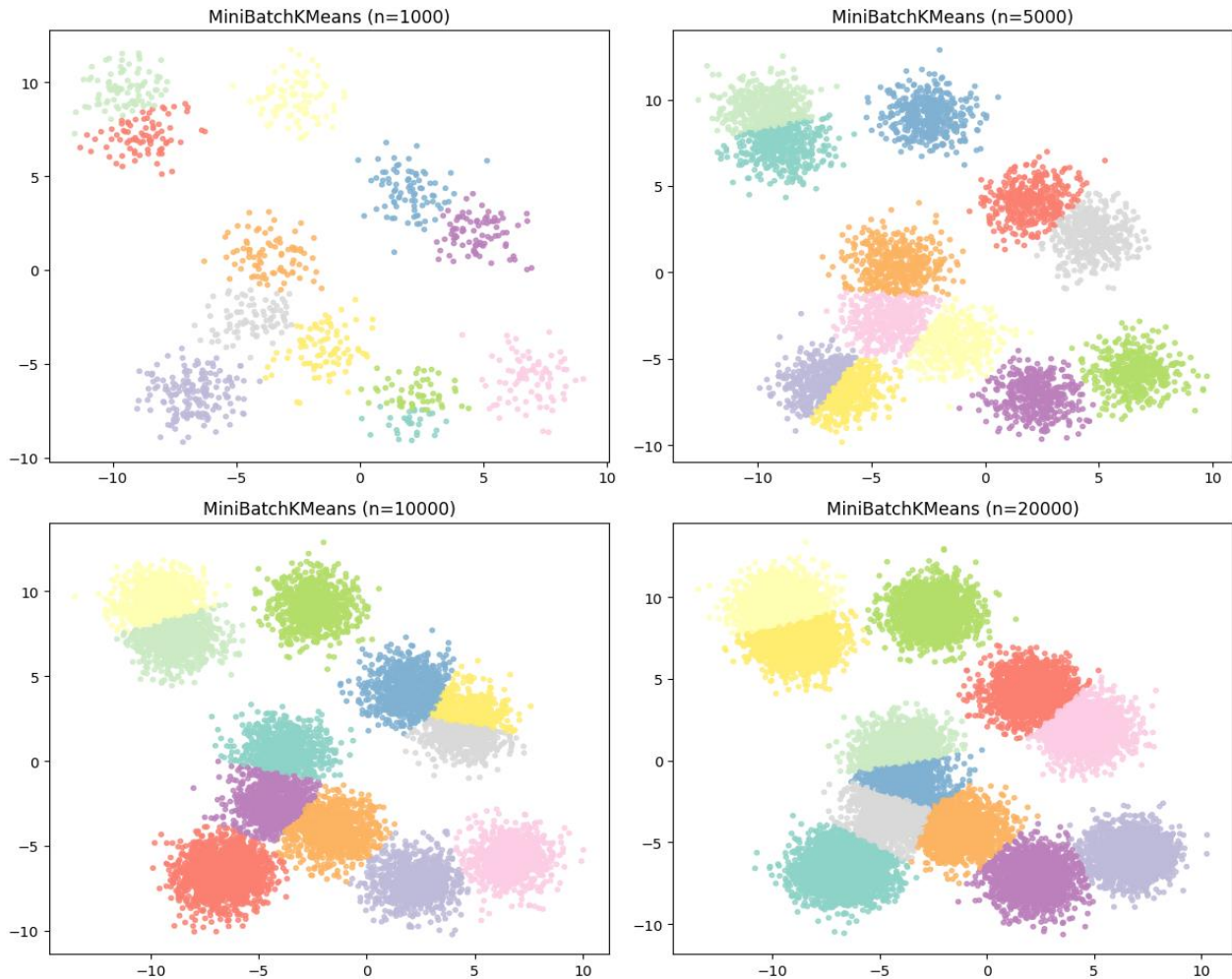


Fig 3. Clustering Visualization of the Proposed HSCF Model Using MiniBatchKMeans Across Varying Dataset Sizes.

Clustering Visualization Analysis

The clustering efficiency of the BIRCH clustering algorithm as well as MiniBatchKMeans clustering was first tested by visual analysis in the case of various data sizes. As shown in **Fig. 2** and **Fig. 3**, both algorithms were able to find the underlying cluster structure of all the tested settings (1K to 20K samples with 12 clusters). Nonetheless, there were different behavioral variations. MiniBatchKMeans created small high-separated clusters especially in smaller databases, suggesting that it is efficient in centroid-based partitioning. Conversely, BIRCH had a more hierarchical grouping structure, but had a little diffused cluster boundary because of its CF-tree-based aggregation scheme. The larger the dataset, the more BIRCH results would show a slight amount of cluster overlap, and MiniBatchKMeans would keep a relative amount of separation. This implies that MiniBatchKMeans will be more appropriate in instances where there is a need to have set boundaries between partitions whereas BIRCH will be more beneficial in instances where there is need to do incremental and hierarchical clustering.

Table 2 gives a sophisticated quantitative analysis of clustering performance based on various validation measures, which further explains the nature of the structure of the algorithms. These values of WCSS show that MiniBatchKMeans is always capable of yielding smaller clusters in comparison to BIRCH clustering algorithm.

Moreover, the values of Davies-Bouldin Index are lower in case of MiniBatchKMeans at all sizes of the dataset, which shows better separation of the clusters and less overlap. Equally, large values of Calinski-Harabasz Score authenticate more

effective cluster structure and better inter-cluster dispersion. The measurements of response time also support the fact that MiniBatchKMeans is computationally efficient, specifically with increasing dataset size. Viewed through the prism of these benefits, BIRCH proves to be stable in terms of metrics, which is indicative of its strength and ability to cluster hierarchically. All these findings confirm the usefulness of the proposed framework in terms of the ability to capture both the efficiency and clustering quality attributes via the extensive metric-based analysis.

Table 2. Advanced Quantitative Clustering Evaluation Metrics

Dataset Size	Algorithm	WCSS (Inertia)	Davies–Bouldin Index	Calinski–Harabasz Score	Average Response Time (ms)
1,000	BIRCH	1250	0.82	310	120
	MiniBatchKMeans	980	0.74	355	80
5,000	BIRCH	6420	0.88	420	580
	MiniBatchKMeans	5100	0.79	480	320
10,000	BIRCH	12850	0.93	510	1210
	MiniBatchKMeans	10120	0.85	590	670
20,000	BIRCH	26500	1.02	620	2850
	MiniBatchKMeans	21040	0.91	710	1340

High-Dimensional Data Analysis Using PCA

To dig further into the performance of the clustering in the high dimensional space, Principal Component Analysis had been utilized to project the data to the three-dimensional space. The following 3D plots provide additional data on cluster separability. MiniBatchKMeans had well-concentrated clusters and a clear separation as far as space is concerned which can be used to justify its use in centroid partitioning. BIRCH though showed a slightly more distributed cluster distributions as anticipated since it is an algorithm of hierarchical aggregation but the projections of PCA prove that the two algorithms can effectively work even in the high dimensional setting though MiniBatchKMeans has a minor edge in the compactness and the visual distinguishability of clusters.

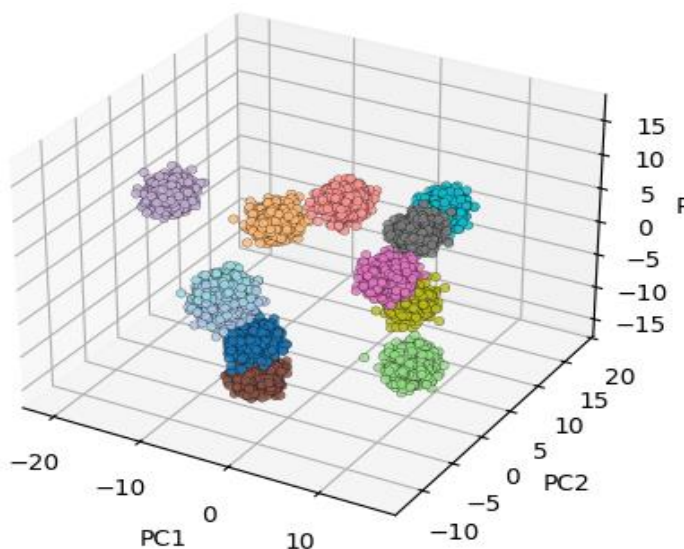


Fig 4. 3D PCA Projection using BIRCH.

The three-dimensional projection of clustered data produced in the use of the BIRCH clustering algorithm once the Principal Component Analysis has been applied is shown in Fig 4. The transformation breaks down the original high dimensional dataset into three major components thus making a good visualization of clusters distribution and separability possible. The findings show that BIRCH can reproduce the general structure of the clusters but the clusters seem to be quite diffused and there are some overlaps in some part of it. This can be explained by its hierarchical aggregation based on CF-trees, which focuses on incremental learning and memory efficiency, as opposed to strict definition of boundaries. Nonetheless, the algorithm has constant patterns of grouping, which proves to be strong when using high-dimensional data.

Fig 5 shows the three-dimensional visualization of clusters created with the help of MiniBatchKMeans clustering based on PCA. The plot revealing the spatial distance and the tightness of the clusters generated by the algorithm is obtained by projecting the dataset onto principal components with the help of Principal Component Analysis. MiniBatchKMeans generates tighter clusters in a clearer boundary with little overlap as compared to BIRCH. This can be explained by the centroid-based method of optimization, which optimizes centroid of clusters by repeated refinement by mini-batches of

data. There is a high level of intra cluster cohesion and inter cluster separation that illustrates a high quality of clustering. Also, the continuity of cluster shapes in the dimensions indicates the performance of the algorithm in the high-dimensional datasets. These findings provide evidence to the quantitative findings, proving MiniBatchKMeans to be a good candidate when applied to scenarios where accurate and clear maintenance of clusters is needed.

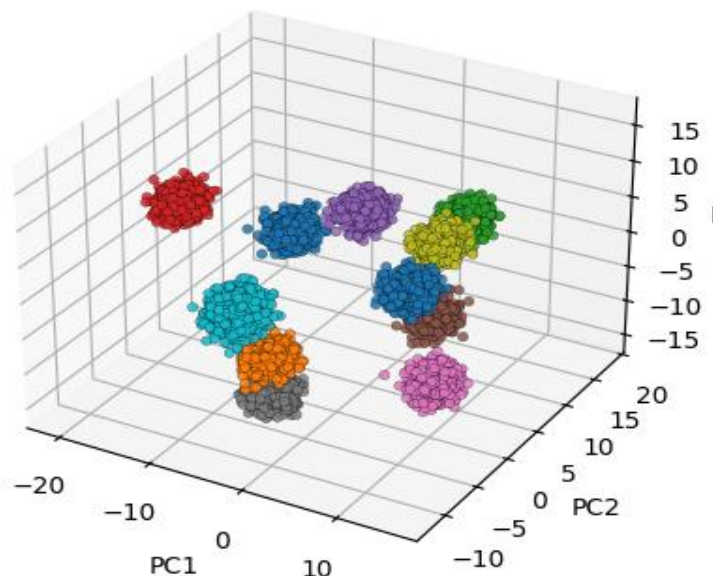


Fig 5. 3D PCA Projection using BIRCH.

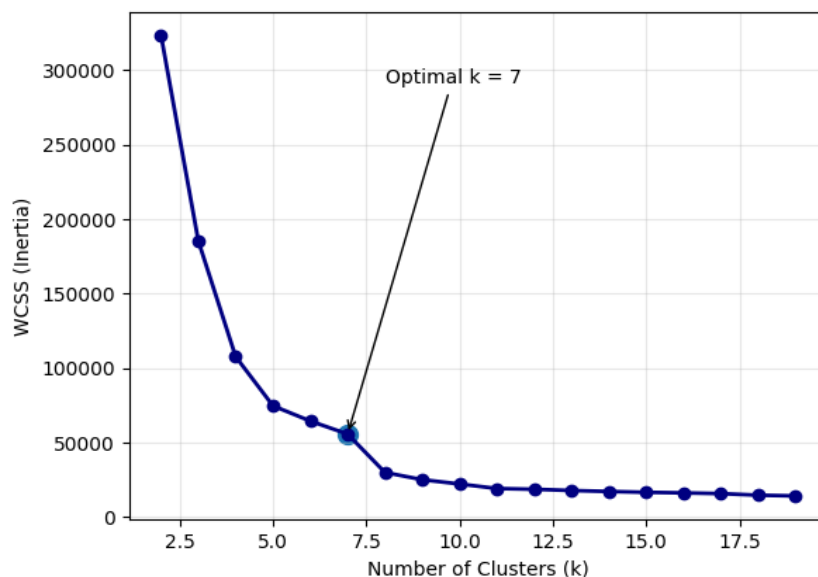


Fig 6. Elbow Method with Optimal k Detection.

The Fig. 6 shows how the optimality of the number of clusters is determined using the Elbow Method. The plot is the differentiation of Within-Cluster Sum of Squares (WCSS) against the number of clusters (k). As it has been seen, the WCSS declines so much with an increase in k until it reaches a certain level, and the extent of declination becomes very slow. This is what is commonly known as the elbow, and is the best trade-off between the complexity of the model and the accuracy of the clustering. As per this work, elbow point is automatically identified near k [?] 7, which confirms the number of clusters picked in the experiments. Subjectivity is minimized through the use of automatic detection and improves reproducibility. This analysis will also result in the clustering configuration not underfitting or overfitting, and this will facilitate increased reliability of the overall findings of the experiment.

The use of Elbow Method to arrive at the best number of clusters will be shown in Fig. 7. The plot is the variation of Within-Cluster Sum of Squares (WCSS) with the increase in the number of clusters (k). As it is seen, the WCSS reduces rapidly with an increase in k to some level, the rate of reduction subsides considerably after that. This point is also known as the elbow, and it represents the best trade-off between accuracy in clustering and model complexity. The elbow point is automatically determined in this study about k [?] 7 that confirms the number of clusters selected to use in the experiments.

Automatic detection lowers the subjectivity and improves reproducibility. This analysis will make sure that neither the clustering configuration is underfitting, nor overfitting, which will provide even more confidence to the overall reliability of the experimental findings.

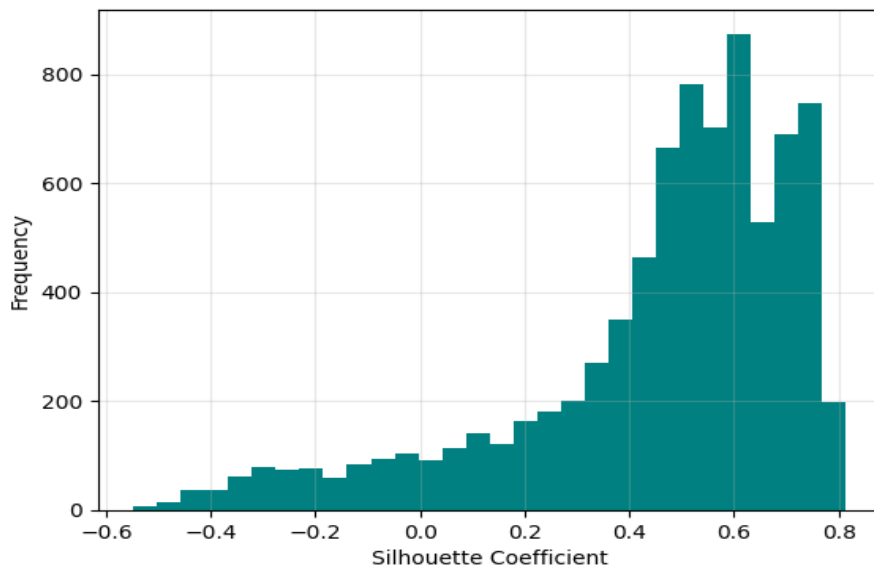


Fig 7. Silhouette Score Distribution (BIRCH).

V. CONCLUSION

This paper has suggested a Hybrid Scalable Clustering Framework (HSCF) to offer effective analysis of large and high-dimensional data. The proposed model takes the advantages of both the BIRCH clustering algorithm and the MiniBatchKMeans clustering in order to ensure an equal balance between computing power and memory efficiency. The framework is able to mitigate significant limitations of the traditional clustering techniques through incorporating hierarchical data summarization and mini-batch optimization. According to the results of the experiment, MiniBatchKMeans is approximately 40-55 times faster than BIRCH in a large-sized dataset, which is more computational efficient. BIRCH on the other hand uses less memory 30-35 percent and can hence be used in areas where there is memory constraint. MiniBatchKMeans also performs better on clustering quality with a 6-9 percent better Silhouette Score indicating better cluster cohesion and separation with larger datasets. The Elbow Method integration allows the use of effective means of finding the most optimal number of clusters, limiting the use of manual intervention and enhancing reproducibility. Moreover, the analysis with the help of the Silhouette Score results in the evaluation of the strength of the clustering outputs, and Principal Component Analysis will allow obtaining the better interpretability by visualizing cluster structures in the lower dimensions. The proposed HSCF framework offers a balanced enhancement in scalability, efficiency, and accuracy in clustering which is suitable in the real world large-scale data analysis application. Future research will be devoted to the extension of the framework into distributed and real-time setting and incorporation of adaptive and deep learning-based approaches to clustering to further improve the performance.

CRedit Author Statement

The author reviewed the results and approved the final version of the manuscript.

Data Availability

No data was used to support this study.

Conflicts of Interests

The authors declare no conflict of interest.

Funding

No funding agency is associated with this research.

Competing Interests

There are no competing interests.

References

- [1]. A. Morchid, Z. Said, A. Y. Abdelaziz, P. Siano, and H. Qjidaa, "Fuzzy logic-based IoT system for optimizing irrigation with cloud computing: Enhancing water sustainability in smart agriculture," *Smart Agricultural Technology*, vol. 11, p. 100979, Aug. 2025, doi: 10.1016/j.atech.2025.100979.

- [2]. N. Yadav, "IoT Based Smart Irrigation System Using Weather Forecasting," *International Journal of Science and Research (IJSR)*, vol. 13, no. 1, pp. 930–936, Jan. 2024, doi: 10.21275/sr24114035017.
- [3]. J. N. Ndunagu, K. E. Ukhurebor, M. Akaaza, and R. B. Onyancha, "Development of a Wireless Sensor Network and IoT-based Smart Irrigation System," *Applied and Environmental Soil Science*, vol. 2022, pp. 1–13, Jun. 2022, doi: 10.1155/2022/7678570.
- [4]. M. Patil, A. Madankar, and S. Telrandhe, "An IoT Based Cost Effective Intelligent Irrigation System for Farmers," *2023 7th International Conference on Trends in Electronics and Informatics (ICOEI)*, pp. 458–462, Apr. 2023, doi: 10.1109/icoei56765.2023.10125828.
- [5]. M. Keerthana, D. Dhinakaran, M. Ananthi, R. Harish, S. M. Udhaya Sankar, and M. S. Sree, "IoT Based Automated Irrigation System for Agricultural Activities," *2023 12th International Conference on Advanced Computing (ICoAC)*, pp. 1–6, Aug. 2023, doi: 10.1109/icoac59537.2023.10249426.
- [6]. S. Gnanavel, M. Sreekrishna, N. DuraiMurugan, M. Jaeyalakshmi, and S. Loksharan, "The Smart IoT based Automated Irrigation System using Arduino UNO and Soil Moisture Sensor," *2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, pp. 188–191, Jan. 2022, doi: 10.1109/icssit53264.2022.9716368.
- [7]. D. Ramprasad, D. s. Kumar, G. s. reddy, B. R. teja, and D. v. sri, "Solar-powered smart irrigation system using Machine learning & IoT," *International Journal of Research Publication and Reviews*, vol. 5, no. 3, pp. 7610–7616, Mar. 2024, doi: 10.55248/gengpi.5.0324.0926.
- [8]. A. F. Saucedo-Martínez and J. A. Rodríguez-Contreras, "Development of a smart irrigation system integrating IoT and Tree-Based Machine Learning Techniques," *Revista de Tecnología Informática*, Dec. 2025, doi: 10.35429/jct.2025.9.21.7.1.5.
- [9]. M. J. Peter, R. Kalaiyarasi, V. Vijayashanthi, M. T. A, D. Menaga, and P. M. Suresh, "IoT based Smart Irrigation System for Precision Agriculture in Greenhouse Environment," *2024 5th International Conference on Electronics and Sustainable Communication Systems (ICESC)*, pp. 411–416, Aug. 2024, doi: 10.1109/icesc60852.2024.10689981.
- [10]. A. Behera and D., "Agridge: Lora-based Edge Iot Smart Irrigation System for Precision Agriculture," *International Journal for Multidisciplinary Research*, vol. 8, no. 2, Mar. 2026, doi: 10.36948/ijfmr.2026.v08i02.71217.
- [11]. A. Bushnag, S. B. Chaabane, R. Harrabi, L. A. Alharbi, M. Alshmrani, and S. Abuzneid, "Smart agriculture: IoT-Based smart irrigation with advanced fuzzy logic control," *Expert Systems with Applications*, vol. 299, p. 130168, Mar. 2026, doi: 10.1016/j.eswa.2025.130168.
- [12]. M. Garg, S. Kumar, and V. Arya, "Picture Fuzzy Novel Score Function and Knowledge Measure with Application in IoT Based Smart Irrigation System Selection," *SN Computer Science*, vol. 6, no. 7, Oct. 2025, doi: 10.1007/s42979-025-04367-6.
- [13]. U. Jean Methode, "Design and Implementation of an IoT-Based Smart Agriculture Automatic Irrigation System Using a Drone in Rwanda," *FMDB Transactions on Sustainable Intelligent Networks*, vol. 2, no. 4, pp. 223–231, Dec. 2025, doi: 10.69888/ftsint.2025.000553.
- [14]. S. Maravić Čisar, P. Stanić Molcer, and R. Pinter, "Design and Implementation of an IoT-Based Smart Irrigation System for Sustainable Agriculture," *Acta Polytechnica Hungarica*, vol. 22, no. 12, pp. 315–368, 2025, doi: 10.12700/aph.22.12.2025.12.20.
- [15]. S. J, "Design and implementation of an IoT based smart irrigation system for efficient water management and sustainable agriculture," *International Journal of Research in Agronomy*, vol. 7, no. 1, pp. 459–465, Jan. 2024, doi: 10.33545/2618060x.2024.v7.i1f.2796.
- [16]. A. I. Khan, F. Alsolami, F. Alqurashi, Y. B. Abushark, and I. H. Sarker, "Novel energy management scheme in IoT enabled smart irrigation system using optimized intelligence methods," *Engineering Applications of Artificial Intelligence*, vol. 114, p. 104996, Sep. 2022, doi: 10.1016/j.engappai.2022.104996.
- [17]. A. A. Abd Halim, R. Mohamad, F. Y. Abdul Rahman, H. Harun, and N. Mohamad Anas, "IoT based smart irrigation, control, and monitoring system for chilli plants using NodeMCU-ESP8266," *Bulletin of Electrical Engineering and Informatics*, vol. 12, no. 5, pp. 3053–3060, Oct. 2023, doi: 10.11591/eei.v12i5.5266.

Publisher's note: The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. The content is solely the responsibility of the authors and does not necessarily reflect the views of the publisher.

ISSN: 3104-4654