Scopus

Documents

Alam Mallik, M.[a b] , Fariza Zulkurnain, N.[a] , Siddiqui, S.[c] , Sarkar, R.[d]

The Parallel Fuzzy C-Median Clustering Algorithm Using Spark for the Big Data
(2024) *IEEE Access*, 12, pp. 151785-151804.

DOI: 10.1109/ACCESS.2024.3463712

[a] International Islamic University Malaysia, Department of Electrical and Computer Engineering, Kuala Lumpur, 53100, Malaysia
[b] Lords Institute of Engineering and Technology, Department of Computer Science and Engineering, Hyderabad, 500091, India
[c] Deccan College of Engineering and Technology, Department of Computer Science and Engineering, Hyderabad, 500001, India
[d] The Assam Royal Global University, Department of Computer Science and Engineering, Assam781035, India

Abstract
Big data for sustainable development is a global issue due to the explosive growth of data and according to the forecasting of International Data Corporation(IDC), the amount of data in the world will double every 18 months, and the Global Data-sphere is expected to more than double in size from 2022 to 2026. The analysis, processing, and storing of big data is a challenging research concern due to data imperfection, massive data size, computational difficulty, and lengthy evaluation time. Clustering is a fundamental technique in data analysis and data mining, and it becomes particularly challenging when dealing with big data due to the sheer volume, velocity, and variety of the data. Big Data frameworks like Hadoop MapReduce and Spark are potent tools that provide an effective way to analyze huge datasets that are being processed by the Hadoop cluster. Apache Spark is one of the most widely used large-scale data processing engines due to its speed, low latency in-memory computing, and powerful analytics. Therefore, we develop a Parallel Fuzzy C-Median Clustering Algorithm Using Spark for Big Data that can handle large datasets while maintaining high accuracy and scalability. The algorithm employs a distance-based clustering approach to determine the similarity between data points and group them in combination with sampling and partitioning techniques. In the sampling phase, a representative subset of the dataset is selected. In the partitioning phase, the data is partitioned into smaller subsets that can be clustered in parallel across multiple nodes. The suggested method, implemented in the Databricks cloud platform provides high clustering accuracy, as measured by clustering evaluation metrics such as the silhouette coefficient, cost function, partition index, clustering entropy. The experimental results show that c=5, which is consistent for cost function with the ideal silhouette coefficient of 1, is the optimal number of clusters for this dataset. A comparative study is done to validate the proposed algorithm by implementing the other contemporary algorithms for the same dataset. The comparison analysis exhibits that our suggested approach outperforms the others, especially for computational time. The developed approach is benchmarked with the existing methods such as MiniBatchKmeans, AffinityPropagation, SpectralClustering, Ward, OPTICS, and BRICH in terms of silhouette index and cost function.

Author Keywords
big data framework; Data clustering; fuzzy C means; fuzzy C median; spark

Index Keywords
Big data, Data assimilation; Big data framework, C-means, Clusterings, Cost-function, Data clustering, Data framework, Explosive growth, Fuzzy C mean, Fuzzy C median, Global issues; Data accuracy

References

- Reinsel, D., Gantz, J., Rydning, J.
(2010) *Data Age 2025: The Evolution of Data to Life-Critical*,
Accessed: Oct. 3

- Ikegwu, A.C., Nweke, H.F., Anikwe, C.V., Alo, U.R., Okonkwo, O.R.
Big data analytics for data-driven industry: A review of data sources, tools, challenges, solutions, and research directions
(2022) *Cluster Comput.*, 25 (5), pp. 3343-3387.
Oct.

- Saeed, M.M., Al Aghbari, Z., Alsharidah, M.
Big data clustering techniques based on spark: A literature review
(2020) *PeerJ Comput. Sci.*, 6, p. e321.
Nov.

- Cooley, R., Mobasher, B., Srivastava, J.
Web mining: Information and pattern discovery on the world wide web
(1997) *Proc. 9th IEEE Int. Conf. Tools Artif. Intell.*, pp. 558-567.
Nov.

- Ansari, Z., Azeem, M.F., Babu, A.V., Ahmed, W.
A fuzzy approach for feature evaluation and dimensionality reduction to improve the quality of web usage mining results
(2012) *Int. J. Adv. Sci., Eng. Inf. Technol.*, 2 (6), pp. 67-73.

- Castellano, G., Mesto, F., Minunno, M., Torsello, M.A.
Web user profiling using fuzzy clustering
(2007) *Applications of Fuzzy Sets Theory (Lecture Notes in Computer Science)*, 4578, pp. 94-101.
F. Masulli, S. Mitra, and G. Pasi, Eds., Berlin, Germany: Springer

- Nasraoui, O., Frigui, H., Krishnapuram, R., Joshi, A.
Extracting web user profiles using relational competitive fuzzy clustering
(2000) *Int. J. Artif. Intell. Tools*, 9 (4), pp. 509-526.
Dec.

- Sarker, I.H.
Machine learning: Algorithms, real-world applications and research directions
(2021) *Social Netw. Comput. Sci.*, 2 (3), p. 160.
May

- Bendechache, M., Tari, A.-K., Kechadi, M.-T.
Parallel and distributed clustering framework for big spatial data mining
(2018) *Int. J. Parallel, Emergent Distrib. Syst.*, 34 (6), pp. 671-689.
Mar.

- Ansari, Z., Babuy, A.V., Ahmed, W., Azeemz, M.F.
A fuzzy set theoretic approach to discover user sessions from web navigational data
(2011) *Proc. IEEE Recent Adv. Intell. Comput. Syst.*, pp. 879-884.
Sep.

- Khalid, M., Yousaf, M.M.
A comparative analysis of big data frameworks: An adoption perspective
(2021) *Appl. Sci.*, 11 (22), p. 11033.
Nov.

- Milosevic, Z., Chen, W., Berry, A., Rabhi, F.A., Buyya, R., Calheiros, R.N., Dastjerdi, A.V.
Real-time analytics
(2016) *Big Data, Princ. Paradigms*, 2016, pp. 39-61.
Jan.

- Saeed, N., Husamaldin, L.
Big data characteristics (V's) in industry
(2021) *Iraqi J. Ind. Res.*, 8 (1), pp. 1-9.
Jun.

- Kumar, S., Mohbey, K.K.
A review on big data based parallel and distributed approaches of pattern mining
(2022) *J. King Saud Univ.-Comput. Inf. Sci.*, 34 (5), pp. 1639-1662.
May

- Neumann, P., Kunkel, J.
High-performance techniques for big data processing
(2020) *Knowledge Discovery in Big Data From Astronomy and Earth*

Observation., pp. 137-158.
Finland: Springer

- Han, J., Kamber, M., Pei, J.
(2012) *Data Mining Concepts and Techniques*,
3rd ed. Waltham, MA, USA: Morgan Kaufmann Publishers

- Chunqiong, W.U.
Research on clustering algorithm based on big data
(2019) *IOP Conf. Ser., J. Phys., Conf. Ser.*, 1237.
Jan.

- Castellano, G., Fanelli, A.M., Torsello, M.A.
Mining usage profiles from access data using fuzzy clustering
(2006) *Proc. 6th WSEAS Int. Conf. Simul., Model. Optim.*, pp. 157-160.

- Mujeeb, P.S.M., PraveenSam, R., Madhavi, K.
An inception towards better big data clustering technique
(2019) *Int. J. Eng. Adv. Technol. (IJEAT)*, 8958, pp. 2249-8958.
Jan.

- Chen, K.-H., Chen, H.-Y., Wang, C.-M.
Bucket MapReduce: Relieving the disk I/O intensity of data-intensive applications in MapReduce frameworks
(2021) *Proc. 29th Euromicro Int. Conf. Parallel, Distrib. Network-Based Process. (PDP)*, pp. 18-25.
Mar.

- Zahid, N., Limouri, M., Essaid, A.
A new cluster-validity for fuzzy clustering
(1999) *Pattern Recognit.*, 32 (7), pp. 1089-1097.
Jul.

- Pietrzykowski, M.
Comparison of mini-models based on various clustering algorithms
(2020) *Proc. Comput. Sci.*, 176, pp. 3563-3570.
Jan.

- Fukuyama, M.S.Y.
A new method of choosing the number of clustersfor fuzzy c-means method
(1989) *Proc. 5th Fuzzy Syst. Symp.*, pp. 247-250.

- Bezdek, J.C., Ehrlich, R., Full, W.
FCM: The fuzzy c-means clustering algorithm
(1984) *Comput. Geosci.*, 10 (2-3), pp. 191-203.
Jan.

- Fong, S., Deb, S., Yang, X.-S., Zhuang, Y.
Towards enhancement of performance of K-means clustering using nature-inspired optimization algorithms
(2014) *Sci. World J.*, 2014, pp. 1-16.
Jan.

- Ansari, Z., Faizabadi, A.R., Afzal, A.
Fuzzy c-least medians clustering for discovery of web access patterns from web user sessions data
(2017) *Intell. Data Anal.*, 21 (3), pp. 553-575.
Jun.

- Ahmed, N., Barczak, A.L.C., Susnjak, T., Rashid, M.A.
A comprehensive performance analysis of apache Hadoop and apache spark for

large scale data sets using HiBench
(2020) *J. Big Data*, 7 (1), p. 110.
Dec.

- Liu, Y., Zhang, X., Chen, J., Chao, H.
A validity index for fuzzy clustering based on bipartite modularity
(2019) *J. Electr. Comput. Eng.*, 2019, pp. 1-9.
Aug.

- Theodoridis, S., Koutroumbas, K.
Cluster validity
(2016) *Pattern Recognit.*, 2, pp. 863-913.
Dec.

- Hewawasam, K.R.G., Premaratne, K., Shyu, M.L.
Rule mining and classification in a situation assessment application: A belief-theoretic approach for handling data imperfections
(2007) *IEEE Trans. Syst., Man, Cybern., B*, 37 (6), pp. 1446-1459.
Dec.

- Liang, S., Han, D., Yang, Y.
Cluster validity index for irregular clustering results
(2020) *Appl. Soft Comput.*, 95.
Oct.

- Yuan, C., Yang, H.
Research on K-value selection method of K-means clustering algorithm
(2019), 2 (2), pp. 226-235.
Jun.

- Fränti, P., Sieranoja, S.
How much can k-means be improved by using better initialization and repeats?
(2019) *Pattern Recognit.*, 93, pp. 95-112.
Sep.

- Zhang, J., Wu, G., Hu, X., Li, S., Hao, S.
A parallel clustering algorithm with MPI-MKmeans
(2013) *J. Comput.*, 8 (1), pp. 10-17.
Jan.

- Sreedhar, C., Kasiviswanath, N., Reddy, P.C.
Clustering large datasets using K-means modified inter and intra clustering (KM-I2C) in Hadoop
(2017) *J. Big Data*, 4 (1), p. 27.
Dec.

- Khang, T.D., Vuong, N.D., Tran, M.-K., Fowler, M.
Fuzzy cmeans clustering algorithm with multiple fuzzification coefficients
(2020) *Algorithms*, 13 (7), p. 158.
Jun.

- Littau, D., Boley, D.
Clustering very large data sets using a low memory matrix factored representation
(2009) *Comput. Intell.*, 25 (2), pp. 114-135.
May

- Xing, H., Chen, H., Lin, H., Wu, X.
An interval type-2 possibilistic c-means clustering algorithm and its application
(2021) *J. Phys., Conf.*, 2132 (1).
Dec.

- Zhang, S., Duan, C.
  Clustering optimization algorithm for data mining based on artificial intelligence neural network
  (2022) *Wireless Commun. Mobile Comput.*, 2022, pp. 1-16.
  Feb.

- He, Y., Tan, H., Luo, W., Mao, H., Ma, D., Feng, S., Fan, J.
  MRDBSCAN: An efficient parallel density-based clustering algorithm using MapReduce
  (2011) *Proc. IEEE 17th Int. Conf. Parallel Distrib. Syst.*, pp. 473-480.
  Dec.

- Fahim, A.
  A clustering algorithm based on local density of points
  (2017) *Int. J. Modern Educ. Comput. Sci.*, 9 (12), pp. 9-16.
  Dec.

- Akman, O., Comar, T., Hrozencik, D., Gonzales, J.
  Data clustering and self-organizing maps in biology
  (2019) *Algebraic and Combinatorial Computational Biology.*, pp. 351-374.
  Cambridge, MA, USA: Academic

- Li, X.-Y., Guo, L.-J.
  Constructing affinity matrix in spectral clustering based on neighbor propagation
  (2012) *Neurocomputing*, 97, pp. 125-130.
  Nov.

- Schafer, J.B., Frankowski, D., Herlocke, J., Sen, S.
  Collaborative filtering recommender systems
  (2007) *The Adaptive Web.*, pp. 291-324.
  Berlin, Germany: Springer-Verlag

- Mallik, M.A., Zulkurnain, N.F., Ahmed, S.K.J., Nizamuddin, M.K.
  Identification of interested web users using decision tree classifier
  (2021) *Advances in Electrical and Computer Technologies (Lecture Notes in Electrical Engineering)*, pp. 143-156.
  Singapore: Springer

- Dash, S., Shakyawar, S.K., Sharma, M., Kaushik, S.
  Big data in healthcare: Management, analysis and future prospects
  (2019) *J. Big Data*, 6 (1), pp. 1-25.
  Dec.

- Oussous, A., Benjelloun, F.-Z., Ait Lahcen, A., Belfkih, S.
  Big data technologies: A survey
  (2018) *J. King Saud Univ.-Comput. Inf. Sci.*, 30 (4), pp. 431-448.
  Oct.

- Liu, B., He, S., He, D., Zhang, Y., Guizani, M.
  A spark-based parallel fuzzy c-means segmentation algorithm for agricultural image big data
  (2019) *IEEE Access*, 7, pp. 42169-42180.

- Sainju, A.M., Aghajarian, D., Jiang, Z., Prasad, S.
  Parallel grid-based colocation mining algorithms on GPUs for big spatial event data
  (2020) *IEEE Trans. Big Data*, 6 (1), pp. 107-118.
  Mar.

- Akthar, N., Ahamad, M.V., Khan, S.
  Clustering on big data using Hadoop MapReduce
  (2015) *Proc. IEEE Int. Conf. Comput. Intell. Commun. Netw. (CICN)*, pp. 789-795.

- Sinha, A., Jana, P.K.
  A novel K-means based clustering algorithm for big data
  (2016) *Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI)*, pp. 1875-1879.
  Sep.

- HajKacem, M.A.B., N'cir, C.-E.-B., Essoussi, N.
  One-pass MapReduce-based clustering method for mixed large scale data
  (2019) *J. Intell. Inf. Syst.*, 52 (3), pp. 619-636.
  Jun.

- Shafiq, M.O., Torunski, E.
  A parallel K-medoids algorithm for clustering based on MapReduce
  (2016) *Proc. 15th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, pp. 502-507.
  Dec.

- Ben Haj Kacem, M.A., Ben N'cir, C.-E., Essoussi, N.
  MapReducebased k-prototypes clustering method for big data
  (2015) *Proc. IEEE Int. Conf. Data Sci. Adv. Analytics (DSAA)*, pp. 1-7.
  Oct.

- Ludwig, S.A.
  MapReduce-based fuzzy c-means clustering algorithm: Implementation and scalability
  (2015) *Int. J. Mach. Learn. Cybern.*, 6 (6), pp. 923-934.
  Dec.

- Hidri, M.S., Zoghlami, M.A., Ben Ayed, R.
  Speeding up the largescale consensus fuzzy clustering for handling big data
  (2018) *Fuzzy Sets Syst.*, 348, pp. 50-74.
  Oct.

- Zhang, Q., Chen, Z.
  A weighted kernel possibilistic c-means algorithm based on cloud computing for clustering big data
  (2014) *Int. J. Commun. Syst.*, 27 (9), pp. 1378-1391.
  Sep.

- Zhang, Q., Yang, L.T., Chen, Z., Li, P.
  PPHOPCM: Privacypreserving high-order possibilistic c-means algorithm for big data clustering with cloud computing
  (2022) *IEEE Trans. Big Data*, 8 (1), pp. 25-34.
  Feb.

- Hu, R., Dou, W., Liu, J.
  ClubCF: A clustering-based collaborative filtering approach for big data application
  (2014) *IEEE Trans. Emerg. Topics Comput.*, 2 (3), pp. 302-313.
  Sep.

- Subramaniyaswamy, V., Vijayakumar, V., Logesh, R., Indragandhi, V.
  Unstructured data analysis on big data using MapReduce
  (2015) *Proc. Comput. Sci.*, 50, pp. 456-465.
  Jan.

- Castellano, G., Fanelli, A.M., Mencar, C., Torsello, M.A.
  Similaritybased fuzzy clustering for user profiling
  (2007) *Proc. IEEE/WIC/ACM Int. Conf. Web Intell. Intell. Agent Technol.-Workshops*, pp. 75-78.
  Nov.

- Sachar, P., Khullar, V.
  Social media generated big data clustering using genetic algorithm
  (2017) *Proc. Int. Conf. Comput. Commun. Informat. (ICCCI)*, pp. 1-6.
  Jan.

- Karimov, J., Ozbayoglu, M.
  High quality clustering of big data and solving empty-clustering problem with an evolutionary hybrid algorithm
  (2015) *Proc. IEEE Int. Conf. Big Data*, pp. 1473-1478.
  Oct.

- Yang, Y., Teng, F., Li, T., Wang, H., Wang, H., Zhang, Q.
  Parallel semisupervised multi-ant colonies clustering ensemble based on MapReduce methodology
  (2018) *IEEE Trans. Cloud Comput.*, 6 (3), pp. 857-867.
  Jul.

- Chen, M., Ludwig, S.A., Li, K.
  Clustering in big data
  (2017) *Big Data: Management, Architecture, and Processing.*, pp. 246-331.
  Boca Raton, FL, USA: CRC Press, ch. 16

- Ezugwu, A.E., Ikotun, A.M., Oyelade, O.O., Abualigah, L., Agushaka, J.O., Eke, C.I., Akinyelu, A.A.
  A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects
  (2022) *Eng. Appl. Artif. Intell.*, 110.
  Apr.

- Dwivedi, Y.K., Ismagilova, E., Hughes, D.L., Carlson, J., Filieri, R., Jacobson, J., Jain, V., Wang, Y.
  Setting the future of digital and social media marketing research: Perspectives and research propositions
  (2021) *Int. J. Inf. Manage.*, 59.
  Aug.

- Nasraoui, O., Krisnapuram, R.
  Clustering using a genetic fuzzy least median of squares algorithm
  (1997) *Proc. Annu. Meeting North Amer. Fuzzy Inf. Process. Soc.*, 3, pp. 217-221.

- Shaikh, E., Mohiuddin, I., Alufaisan, Y., Nahvi, I.
  Apache spark: A big data processing engine
  (2019) *Proc. 2nd IEEE Middle East North Afr. Commun. Conf. (MENACOMM)*, pp. 1-6.
  Nov.

- Mavai, A.S., Mishra, S.K.
  A survey and comparative study of different data mining techniques to implement a missing value estimator system
  (2014) *Int. J. Current Eng. Technol.*, 4 (4).
  Aug.

- Ansari, Z., Azeem, M.F., Babu, A.V., Ahmed, W.
  A fuzzy clustering based approach for mining usage profiles from web log data
  (2011) *Int. J. Comput. Sci. Inf. Secur.*, 9, pp. 70-79.
  Jun.

- Peters, G., Crespo, F., Lingras, P., Weber, R.
  Soft clustering-Fuzzy and rough approaches and their extensions and derivatives
  (2013) *Int. J. Approx. Reasoning*, 54 (2), pp. 307-322.
  Feb.

- Zadeh, L.A.
  Recent developments and new directions in soft computing
  (2014) *Studies in Fuzziness and Soft Computing*, 317.
  Berlin, Germany: Springer

- Davara, V., Upadhyay, A.B.
  Comparison of soft computing techniques for the design of microstrip patch antenna: A review paper
  (2014) *Int. J. Adv. Res. Electr., Electron. Instrum. Eng.*, 3.
  Mar.

- Xie, X.L., Beni, G.
  A validity measure for fuzzy clustering
  (1991) *IEEE Trans. Pattern Anal. Mach. Intell.*, 13 (8), pp. 841-847.
  Aug.

- Ikotun, A.M., Ezugwu, A.E.
  Boosting k-means clustering with symbiotic organisms search for automatic clustering problems
  (2022) *PLoS ONE*, 17 (8).
  Aug.

- Zanaty, E.A.
  Determining the number of clusters for kernelized fuzzy C-means algorithms for automatic medical image segmentation
  (2012) *Egyptian Informat. J.*, 13 (1), pp. 39-58.
  Mar.

- Rajkumar, K.V., Yesubabu, A., Subrahmanyam, K.
  Fuzzy clustering and fuzzy c-means partition cluster analysis and validation studies on a subset of citescore dataset
  (2019) *Int. J. Electr. Comput. Eng. (IJECE)*, 9 (4), p. 2760.
  Aug.

- Haleem, A., Javaid, M., Qadri, M.A., Singh, R.P., Suman, R.
  Artificial intelligence (AI) applications for marketing: A literaturebased study
  (2022) *Int. J. Intell. Netw.*, 3, pp. 119-132.
  Jan.

- Khalil, W., Torkey, H., Attiya, G.
  Survey of apache spark optimized job scheduling in big data
  (2020) *Int. J. Ind. Sustain. Develop.*, 1 (1), pp. 39-48.
  Jan.

- Armoogum, S., Li, X.
  Big data analytics and deep learning in bioinformatics with Hadoop
  (2019) *Deep Learning and Parallel Computing Environment for Bioengineering Systems.*, pp. 17-36.
  The Netherlands: Elsevier

- Islam, M.T., Srirama, S.N., Karunasekera, S., Buyya, R.
  Costefficient dynamic scheduling of big data applications in apache spark on cloud
  (2020) *J. Syst. Softw.*, 162.
  Apr.

- Petrov, M., Butakov, N., Nasonov, D., Melnik, M.
  Adaptive performance model for dynamic scaling apache spark streaming
  (2018) *Proc. Comput. Sci.*, 136, pp. 109-117.
  Jan.

- Basha, S.A.K., Basha, S.M., Vincent, D.R., Rajput, D.S.
  Challenges in storing and processing big data using Hadoop and spark
  (2019) *Deep Learning and Parallel Computing Environment for Bioengineering Systems*, pp. 179-187.
  The Netherlands: Elsevier

- (2021) *How to Use Spark Clusters for Parallel Processing Big Data*,
  Accessed: Dec. 3

- Ramírez-Gallego, S., García, S., Benítez, J.M., Herrera, F.
  A distributed evolutionary multivariate discretizer for big data processing on apache spark
  (2018) *Swarm Evol. Comput.*, 38, pp. 240-250.
  Feb.

- (2022) *Introduction to Apache Spark With Examples and Use Cases*,
  Accessed: Aug.

- (2022) *PySpark-SparkContext*,
  Accessed: Jul.

- (2022),
  Accessed: Jul.

- (2022),
  Accessed: Jul.

- Zaharia, M., Xin, R.S., Wendell, P., Das, T., Armbrust, M., Dave, A., Meng, X., Ghodsi, A.
  Apache spark: A unified engine for big data processing
  (2016) *Commun. ACM*, 59, pp. 56-65.
  Oct.

- (2020),
  Accessed: Jul. 15

- Verma, A., Mansuri, A.H., Jain, N.
  Big data management processing with Hadoop MapReduce and spark technology: A comparison
  (2016) *Proc. Symp. Colossal Data Anal. Netw. (CDAN)*, 4, pp. 1-4.
  Mar.

- Thiruvathukal, G.K., Christensen, C., Jin, X., Tessier, F., Vishwanath, V.
  (2019) *A benchmarking study to evaluate apache spark on large-scale supercomputers*,

- Gopalani, S., Arora, R.
  Comparing apache spark and map reduce with performance analysis using K-means
  (2015) *Int. J. Comput. Appl.*, 113 (1), pp. 8-11.
  Mar.

- Usman, S., Mehmood, R., Katib, I., Albeshri, A.
  Data locality in high performance computing, big data, and converged systems: An analysis of the cutting edge and a future system architecture
  (2022) *Electronics*, 12 (1), p. 53.
  Dec.

- Zhao, Y., Dong, J., Liu, H., Wu, J., Liu, Y.
  Performance improvement of DAG-aware task scheduling algorithms with efficient cache management in spark
  (2021) *Electronics*, 10 (16), p. 1874.
  Aug.

- Sterling, T., Anderson, M., Brodowicz, M.
  MapReduce
  (2018) *High Performance Computing.*, pp. 579-589.
  The Netherlands: Elsevier

- Aranganayagi, S., Thangavel, K.
  Clustering categorical data using silhouette coefficient as a relocating measure
  (2007) *Proc. Int. Conf. Comput. Intell. Multimedia Appl. (ICCIMA)*, pp. 13-17.
  Dec.

- Paul, A.K., Shill, P.C.
  New automatic fuzzy relational clustering algorithms using multi-objective NSGA-II
  (2018) *Inf. Sci.*, 448-449, pp. 112-133.
  Jun.

- Pérez-Ortega, J., Moreno-Calderón, C.F., Roblero-Aguilar, S.S., Almanza-Ortega, N.N., Frausto-Solís, J., Pazos-Rangel, R., Rodríguez-Lelis, J.M.
  A new criterion for improving convergence of fuzzy C-means clustering
  (2024) *Axioms*, 13 (1), p. 35.
  Jan.

- Zhang, J., Ma, Z.
  Hybrid fuzzy clustering method based on FCM and enhanced logarithmical PSO (ELPSO)
  (2020) *Comput. Intell. Neurosci.*, 2020, pp. 1-12.
  Mar.

- Shi, M., Zhang, T., Zhang, L., Sun, W., Song, X.
  A fuzzy c-means algorithm based on the relationship among attributes of data and its application in tunnel boring machine
  (2020) *Knowl.-Based Syst.*, 191.
  Mar.

- Ren, C., Sun, L.
  A bi-directional fuzzy c-means clustering ensemble algorithm considering local information
  (2021) *Int. J. Comput. Intell. Syst.*, 14 (1), pp. 1-14.
  Dec.

- Razavi, S.M., Kahani, M., Paydar, S.
  Big data fuzzy c-means algorithm based on bee colony optimization using an apache hbase
  (2021) *J. Big Data*, 8 (1), pp. 1-22.
  Dec.

- Al-Asadi, M.A., Tasdemir, S.
  Using artificial intelligence against the phenomenon of fake news: A systematic literature review
  (2022) *Combating Fake News With Computational Intelligence Techniques (Studies in Computational Intelligence)*, 1001.
  M. Lahby, A. S. K. Pathan, Y. Maleh, W. M. S. Yafooz, Eds., Cham, Switzerland: Springer

- Aiden, M.K., Sabharwal, S.M., Chhabra, S., Al-Asadi, M.
  AI and blockchain for cyber security in cyber-physical system
  (2023) *AI Models for Blockchain-Based Intelligent Networks in IoT Systems (Engineering Cyber-Physical Systems and Critical Infrastructures)*, 6.
  B. Bhushan, A. K. Sangaiah, T. N. Nguyen, Eds., Cham, Switzerland: Springer

- Al-Asadi, M.A., Tasdemir, S.
  Medical image analysis using machine learning techniques
  (2022) *Machine Learning and Deep Learning in Efficacy Improvement of Healthcare*

Systems, pp. 137-153.
U. K.: Taylor & Francis

- Kumari, A., Golyan, A., Shah, R., Raval, N.
Introduction to data analytics
(2024) *Recent Trends and Future Direction for Data Analytics*, pp. 1-14.
A. Kumari, Ed., Hershey, PA, USA: IGI Global

Correspondence Address
Alam Mallik M.; International Islamic University Malaysia, Malaysia; email: malammallik@lords.ac.in

Publisher: Institute of Electrical and Electronics Engineers Inc.

ISSN: 21693536
Language of Original Document: English
Abbreviated Source Title: IEEE Access
2-s2.0-85204715909
Document Type: Article
Publication Stage: Final
Source: Scopus