HFMST: An Efficient Adaptive Fuzzy Linkage Feature Selection with Hybrid Fuzzy Based Minimum Spanning Tree (HFMST) Clustering Algorithm

L. Dhanapriya[1]
Research Scholar, Department of Computer Science, Sri Ramakrishna College of Arts and Science for Women, Coimbatore, Tamil Nadu, India

Dr. S. Preetha[2]
Associate Professor, Department of Computer Science, Sri Ramakrishna College of Arts and Science for Women, Coimbatore, Tamil Nadu, India

Abstract: In this paper, the HFMST algorithm—a novel hybrid Fuzzy Based Minimum Spanning Tree (HFMST) clustering—is presented. A new clustering approach known as HFMST clustering algorithm is created by merging Enhanced Adaptive Fuzzy Linkage based Feature Selection (EAFFS) algorithm with MST. The current feature selection approaches have a high likelihood of redundant features showing up in the final subset, but in most circumstances, locating and eliminating them can significantly increase the clustering accuracy. EAFFS and the HFMST methodology are two approaches used to address this issue. The standard clustering centre is upgraded to take into account the sample weight and the cluster centre of the cluster group when determining the position in the cluster tree. Up to the point of the algorithm convergence iteration optimizes the cluster centres and partitions while also using MST to recalculate the cluster centre. Comparing the proposed HFMST algorithm to other clustering algorithms like DBSCAN, DCNaN, and RDMN algorithms reveals that it has a high Rand Index (RI) and Adjusted Rand Index (ARI) ratio.

Keywords: Data mining, Feature Selection, MST, Distance.

I. INTRODUCTION

The most recent interdisciplinary area of computational science is data mining. The practice of extracting valuable information from vast amounts of data that have been stored in databases, data warehouses, or other information repositories is known as data mining. It is a method for automatically identifying data patterns in a sizable database [1]. Clustering in data mining is the assembling of related data elements into groups called clusters. One of the primary analytical techniques used in data mining is cluster analysis, and the clustering algorithm's design will directly affect the clustering outcomes. To efficiently cluster datasets with various features, a variety of clustering approaches have been presented in the literature [2].

Data mining algorithms typically have processing difficulties when dealing with data that has a lot of features or attributes. Model attributes are the parameters of the algorithm's processing area. The computational cost involved in algorithmic processing increases with the dimensionality of the processing space. The majority of Data Mining methods encounter a significant barrier to efficiency in the form of dimensionality. Irrelevant features merely amplify the noise in the data and reduce the model's precision. Noise makes the model larger, requiring more time and system resources to develop and score. Additionally, data sets containing a lot of attributes may include clusters of associated attributes. These characteristics can be assessing the same underlying property.

Dimension reduction in some form is occasionally a desired preprocessing step for data mining in order to reduce the impacts of noise, correlation, and high dimensionality. Dimension reduction is a step in the feature selection process. The need for dimensionality reduction is growing, and as a result, research on feature selection has become deeply and broadly integrated into a variety of domains, including computational statistics, pattern recognition, machine learning, data mining, and knowledge discovery.

Feature selection in data mining is a useful technique for finding pertinent characteristics for dimension reduction. There is no performance degradation when features are eliminated, according to numerous researches. Depending on whether the training data are labeled or not, supervised and unsupervised feature selection algorithms can be created. The two categories of feature selection algorithms have primarily been explored separately up to this point. Unsupervised feature selection, which does not require labels, takes advantage of data variance and separability to assess feature relevance instead of using the correlation between the feature and the class in supervised feature selection. Generally speaking, choosing a feature is a search problem based on some evaluation criteria.

Data is gathered into groups (i.e., clusters) based on correspondence, which is known as clustering [3]. The relationships between patterns within groups are stronger than the relationships between outlines in other clusters. The task of data categorization, on the other hand, is to assign things to one of many predetermined kinds based on training data from previously classified items [4].

It takes a lot of effort to work with high-dimensional data, and accurate information extraction and interpretation require specialized domain knowledge. The weight for each edge in MST-based grouping algorithms is often calculated as the Euclidean spacing between the focuses interacting with that edge. We are able to overcome a substantial fraction of the problems that the conventional grouping algorithms face thanks to Least Spanning Tree (MST) based clustering techniques. MST-based grouping algorithms are commonly used in practice because of their ability to recognize clusters with erratic bounds.

Graph theoretic clustering is connected to the hierarchical clustering techniques. MST is utilized in minimal spanning tree clustering techniques. The rate of clustering can be reduced since the MST overlooks numerous potential links between the data patterns. It is well known that the MST-based clustering technique may find clusters of different sizes and forms [5]. The MST clustering algorithm does not presume that the underlying data is organized into spherical forms, in contrast to conventional clustering algorithms. The EMST clustering technique [6] produces the structure of point clusters in n-dimensional Euclidean space using the Euclidean minimal spanning tree of a graph. Clusters are discovered to achieve specific optimality metrics, such as the lowest intra-cluster distance or the greatest inter-cluster distance [7].

Typically, a clustering technique is employed to partition the data in order to intra-cluster and inter-cluster similarities are maximized and minimized, respectively, the similarities within and between clusters [7]. The related study [8] proposes a wide range of methods to accomplish this objective, with the Relative Density Measure Based on MST Neighborhood [9] clustering. First, MST is built for a given input using this procedure [10].

The goal of this work is to provide a capable data mining technique for feature selection and clustering that works well to assess a density and manage the scale, which affects the clustering results. This paper presents the methodologies namely data preprocessing, Adaptive Fuzzy Linkage based Feature Selection (AFFS) algorithm and Enhanced Hybrid Fuzzy based Minimum spanning tree (HFMST) clustering is used to discover the good cluster patters result and time efficiency.

II. RELATED WORK

Cheng et al. (2016) [11] addressed how most spatial clustering algorithms become useless. The authors suggested a brand-new clustering technique called SCDOT (Spatial Clustering with Density-Ordered Tree) to overcome this problem. After a dataset is projected to a Density-Ordered Tree, SCDOT divides it into a number of little sub-clusters using a box-plot method. The true clusters are found by repeatedly merging sub-clusters, and an iteration mechanism is employed to pick the input parameters automatically.

Thomas et al. (2017) [12] illustrate a cluster's structural traits are depicted in these graphs. The new index is effective on clusters of various sizes and forms, and it does this by overcoming the drawbacks of conventional indices based on statistics measurements. These graphs are produced using an iterative approach based on principal component analysis, which divides the clusters into a set number of "sub-clusters" that can be adjusted. Then, a minimal spanning tree based on the centroids of each of these sub-clusters is constructed and used to calculate the distances between them as well as the quality of the clusters.

Using minimal spanning trees (MSTs) to estimate Rényi mutual information, Eggels, A., and Crommelin, D. (2019) [13] introduced a unique approach for assessing interdependence in multivariate datasets. For purposes like uncertainty quantification and sensitivity analysis, it is crucial to know how dependent random variables are. The latter is directly connected to the question of how heavily dependent the result of, say, a computer simulation is on each particular random input variable.

Sieranoja, S., and Fränti, P., (2019) [14] illustrated to address the temporal complexity issue. For both calculating density and delta, the suggested technique builds an approximate k-nearest neighbor graph quickly and generically. As long as a distance function is given, this method maintains the generality of density peaks, allowing it to be used for all forms of data. Their method delivers a 91:1 speedup factor for a dataset with a size of 100,000.

According to Gagolewski, M. (2021) [15], the method frequently beats other cutting-edge techniques in terms of clustering quality and speed. It also supports a range of distances over dense, sparse, and string data domains, and it can be made even more resilient by using the built-in noise point detector. It can be used to solve issues that arise in all data-driven research and development activities, including those in the environmental technology and engineering, because it is domain-independent software.

Ma et al. 2021 [16] discussed as tends to make the merging process less complex while still guaranteeing clustering performance. The suggested approach is based on two main concepts: (1) the centroid of MST rather than the cluster's centre is used to compute the intercluster distance. (2) The merge criterion is taken to be the length of the cut edge where two consecutive clusters intersect.

Amit Kumar Das et al. 2022 [17] discussed an efficient methods for reducing dimensionality, all of them—aside from feature selection—do not maintain the original features of the input dataset. The importance of features is evaluated, as well as any potential redundancy for dimensionality reduction, during feature selection. The authors proposed a novel fuzzy graph-based method for determining the relevance and redundancy of features.

The establishment of external evaluation measures for soft clustering was the subject of a comprehensive framework provided by Campagner et al. in 2023 [18]. Their plan is predicated on the idea that soft clustering represents hazy information about an underlying, unidentified hard grouping. Using a broad formulation based on optimal transport theory, the authors showed how any evaluation metric can be easily extended to soft clustering. The suggested "transport-based measure" offers an unbiased, interval-valued comparison index that depicts the range of compatibility between two soft clustering's. They investigated the proposed method's metric and complexity characteristics, as well as its compatibility with other ideas that have already been put out.

III. RESEARCH METHODOLOGY

The proposed approach successfully tests every experiment that was carried out on a large-scale real-world dataset. The research methodology considers an efficient clustering process using data preprocessing, Enhanced Adaptive Fuzzy Linkage based Feature Selection (EAFFS) algorithm and Enhanced Hybrid Fuzzy based Minimum spanning tree (HFMST) clustering process are derived in this section. The proposed clustering process considers overall process flow diagram is described in figure 1.

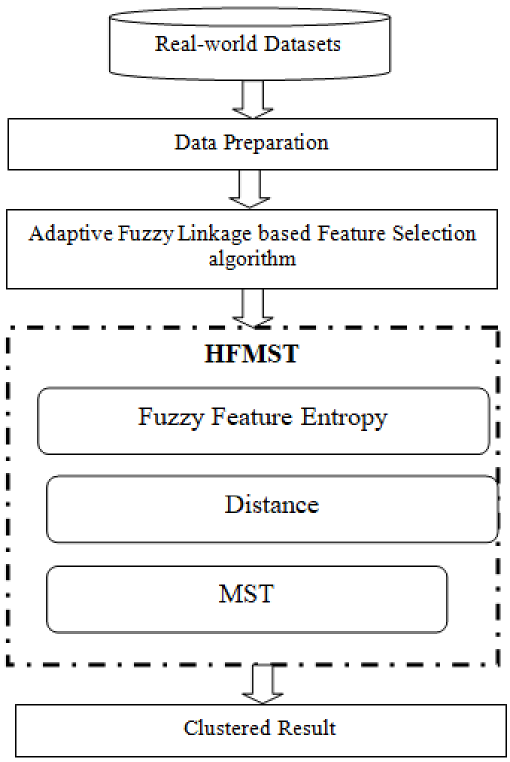

Fig. 1: Proposed Workflow Architecture Process flow

A. *DATA PREPROCESSING*

Data preprocessing is the procedure of data cleaning. The group of perfect data ensures that the connected decisions are accurate. Data is gathered from a variety of data sources found on various web sources. The datasets were acquired from the Gaussian Data Set [19] and the UCI data repository [20], and they include information on including Lukemia, Lung Cancer, and Gaussian distributed dataset. Table 1 describes the real-world datasets.

Table 1: Real-world Dataset

Dataset	Number of Instances	Cluster	Number of Features
Lukemia	248	6	985
Lung Cancer	32	3	56
MiniBooNE	130064	2	50
Dim-1024	1024	16	32
BioTrain	145751	2	74
Novaratis	103	4	1000
LungA	197	4	1000

Data preparation, also known as data cleaning, is a crucial step in the data mining process that involves converting special data properties into an arithmetic format. The preprocessing procedure makes an effort to reduce NaN (Not a Number) data, noise, and inappropriate data that are present in the dataset. Preprocessing contribution data sets for knowledge discovery objective techniques often uses the largest portion of the effort allocated in the complete work in data mining.

B. Enhanced Adaptive Fuzzy Linkage based Feature Selection (EAFFS) algorithm

The Enhanced Adaptive Fuzzy Linkage Based Feature Selection (EAFFS) approach is evolved from the fuzzy clustering technique. To create cluster centers and create a membership function to fuzzify all features, a fuzzy C Means clustering algorithm is utilized. Apply the membership function to each feature in the dataset to determine the fuzzy entropy. The second stage, which focuses on boundary samples, chooses feature subsets based on the suggested fuzzy entropy measure.

Let a data set $D = \{d_k | k=1,2, \ldots, n\}$ incorporate n amount of data's $dk \in R^q$; q is the number of features $feat_j$. To denote $d^k{}_m$ as the significance of the m^{th} feature $feat_m$ of dataset d_k, then every data d_k of set D can be symbolized by a vector

$$D_k = \{d_k^1, d_k^2, \ldots, d_k^q\} \qquad (1)$$

Assume that a set *Feat* represents all features:

$$Feat = \{feat_1, feat_2, \ldots, feat_q\} \qquad (2)$$

To use a membership set U defined as to scheme the initial data set to a fuzzy space

$$U = \{\mu_{11}, \mu_{12}, \ldots, \mu_{1a}, \mu_{21}, \mu_{22}, \ldots, \mu_{2b\ldots,}\mu_{p1\ldots,,}\mu_{pc}\} \qquad (3)$$

where factor μ_{mn} is n^{th} fuzzy group of feature $feat_m$. The optimistic keys a, b, and c describe the cardinality of the fuzzy sets of the first ($feat_1$), second ($feat_2$), and pth feature ($feat_p$), respectively.

The fuzzy estimated group $Feat_D$ of the unique data set D defined as $F_D=\{(d_k, \mu(d_k)| \text{ k}= 1,2,\ldots,\text{n}\}$ where $\mu(d_k)$ is a vector represented as,

$$\mu(d_k) = \left\{ \begin{matrix} \mu_{11}(d_k^1), \mu_{12}(d_k^1), \ldots, \mu_{1a}(d_k^1), \mu_{21}(d_k^2), \mu_{22}(d_k^2) \\ , \ldots, \mu_{2b}(d_k^2) \ldots, \mu_{p1}(d_k^p) \ldots, \mu_{pc}(d_k^p) \end{matrix} \right\} \qquad (4)$$

The membership of each pattern in each cluster is calculated using the fuzzy linkage method (centroids vectors m_j), and the membership of each distinct pattern xk in each cluster is then normalized.

$$cl_c(ft) = \frac{\sum_{d \in Dc} \mu_{ft}(x)}{\sum_{d \in D} \mu_{ft}(x)} \qquad (5)$$

Dc denotes the samples of class c , c ε C , $\mu_{ft}(x)$ denotes the membership grade of x belonging to the fuzzy set ft , $\mu_{ft}(x) \in [1,0]$.

The sum of fuzzy linkage entropy of the instances in feature *feat*,

$$FL\_Entropy(feat) = \sum_{v \in V} \left(\frac{FL_v}{FL}\right) \sum_{c \in C} (-cl_c(v) Log_2 cl_c(v))) \qquad (6)$$

The final adaptive Fuzzy linkage feature measure suing equation 7,

$$FLFM(C, feat) = EC(C) - FL_{Entropy(feat)} \qquad (7)$$

Where EC is entropy class,

$$EC(C) = \sum_{i=1}^{n} p_i log_2 p_i \qquad (8)$$

Algorithm 1: EAFFS
Input: Input Dataset *D*, features *feat*, class *C*
Output: A relevant features
Process
Step 1: Initialize the finest subset of features as the empty set.
Step 2: Calculate fuzzy membership U using eqn. (3)
Step 3: Calculate fuzzy linkage entropy using eqn. (5)
Step 4: Sort the characteristics in order of dependency measure.
Step 4: Select only the best features.

C. ENHANCED HYBRID FUZZY BASED MINIMUM SPANNING TREE (HFMST) CLSUTERING

This paper presents a new method of Enhanced Hybrid Fuzzy based minimum spanning tree (HFMST) clustering is an extension of Fuzzy and MST clustering model. A predetermined cluster number is not necessary for the HFMST algorithm, which is based on the Minimum Spanning Tree. In order to eliminate the inconsistent edges that fulfill the inconsistence measure, the method builds an HFMST out of a point set. Up until the ideal number of clusters is attained, the procedure is constant to build a hierarchy of clusters. Outliers can be quickly identified by using the right number of clusters.

The HFMST approach begins by building a Minimum Spanning Tree (MST) from a point set *Ps* in $FLFM^n$. Euclidean distance between the two end points is what gives an edge in a tree its weight. There is no predetermined cluster number needed for the proposed HFMST approach. The algorithm operates in two steps. The algorithm's first step divided the regions of the fuzzy cluster. Using point deviation, one may locate the centres of clusters or regions. These points are a representative point for the each subtree Ps_T. A feature point f_i is assigned to a cluster i if $c_i \in T_i$. where, disjoint subtrees $Ps_T = \{T_1, T_2, \ldots Tm\}$. Every of these subtrees T_m is delighted as cluster having D data ponts of feature size. The group of center points is represented as $Cnt = \{cnt_1, cnt_2, \ldots, cnt_k\}$. These center points $\{cnt_1, cnt_2, \ldots, cnt_k\}$ are linked and HFMST cluster is constructed.

A related weighted edge can be used to indicate a cluster pair's Euclidean distance. The least spanning tree is the foundation of our algorithm as well, although it is not restricted to two-dimensional points. The algorithm's second step, which uses the ideal number of clusters, is utilized to identify the outliers.

Let D = $\{d_i.\ i = 1,2, \ldots N\}$ be a set of data points (vectors) of feature size N in the subtree (cluster), where (d_{ij}, j = 1, 2, … K) is a K dimensional data point. For a given query point qp = (qp_j= 1,2, ... K), it is required to find the data with minimum distance from the set D. Under the squared-error distance metric, the separation between query point qp and data set D is specified as,

$$Error\_Distnace(qp, d_i) = \sum_{j=1}^{K} (qp_j - d_{ij})^2 \quad (9)$$

IV. Experimental Result

Using MATLAB R2018a simulation, it is seen through experimental assessment that the suggested HFMST algorithm performs well when run on an Intel I7 series 3.21 GHz, x64-based CPU, with 16 GB of central memory, and Windows 10 operating system. To compare the proposed HFMST performance with existing DBSCAN [21], DCNaN[22] and RDMN [9] algorithms was performed. Real-world, low- and high-dimensional datasets are attainable in this experiment. The proposed HFMST technique was used to calculate dataset accuracy and construct the needs for various experiment types. The comparison of Rand Index (RI) measure described in Table 2 and figure 2. The following is the definition of Rand Index [23]:

$$RandIndex\ (RI) = (a + b)/nC_2) \quad (10)$$

Where,

a: The amount of times a pair of elements belongs to the same cluster.

b: The amount of times a pair of elements belong to difference clusters.

nC_2: The number of unordered pairs in a group of n elements.

Table 2: Comparison of RI Values for Different Techniques on Real Datasets

Dataset	DBSCAN	DCNaN	RDMN	**Proposed HFMST**
Lukemia	0.3305	0.7627	0.5539	0.9264
Lung Cancer	0.5060	0.3468	0.5926	0.7241
MiniBooNE	0.5962	0.5962	0.6033	0.6841
Dim-1024	0.8670	0.9255	0.9547	0.9708
BioTrain	0.7172	0.9793	0.9823	0.9912
Novaratis	0.4083	0.6457	0.6587	0.7185
LungA	0.5482	0.6727	0.7727	0.8154

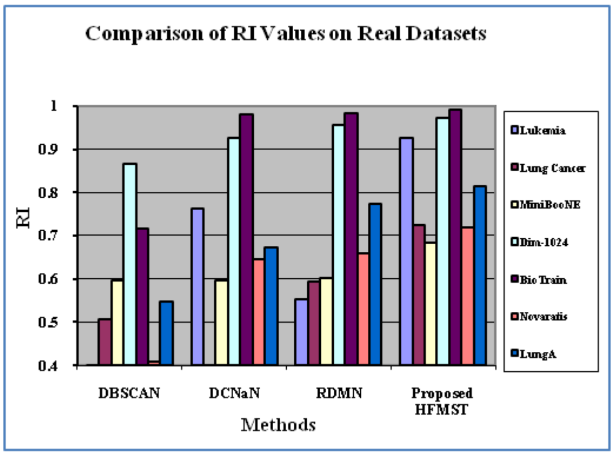

Fig. 2: RI chart

The comparison of Adjusted Rand Index (ARI) measure described in Table 3 and figure 3.

Table 3: Comparison of Adjusted Rand Index (ARI) Values on Real Datasets

Dataset	DBSCAN	DCNaN	RDMN	Proposed HFMST
Lukemia	0.3255	0.4011	0.6311	0.7513
Lung Cancer	0.2344	0.1258	0.2288	0.3842
MiniBooNE	0.147	0.258	0.332	0.4235
Dim-1024	0.2488	0.3111	0.4288	0.4821
BioTrain	0.23585	0.3841	0.4299	0.5233
Novaratis	0.3192	0.3130	0.3300	0.3841
LungA	0.1270	0.3615	0.5344	0.6018

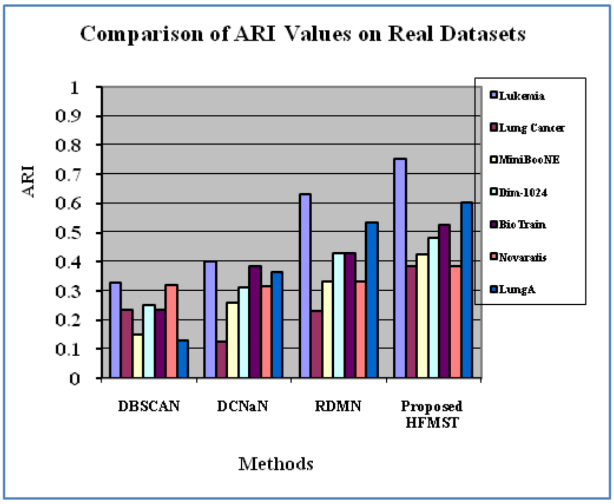

Fig. 3: ARI chart

V. CONCLSUION

This paper presents a enhanced feature selection process with Hybrid Fuzzy Based Minimum Spanning Tree (HFMST) clustering technique is analyzed. The proposed HFMST algorithm is implemented with three stages namely Data preprocessing, Enhanced Adaptive Fuzzy Linkage based Feature Selection (EAFFS) algorithm, and HFMST Clustering. Data cleaning is used to perform preprocessing on the real-world low and high dimensional datasets. The EAFFS approach is used to remove the irrelevant features with error distance measure weight from the proposed feature selection. The proposed HFMST clustering algorithm combines Fuzzy with MST algorithm to automatically partition feature samples based on the obtained cluster.

REFERENCES

[1] S. Sharma, J. Agrawal, S. Agarwal, and S. Sharma, “Machine learning techniques for data mining: A survey,” 2013 IEEE Int. Conf. Comput. Intell. Comput. Res. IEEE ICCIC 2013, no. I, 2013.

[2] A. K. Jain, “Data clustering: 50 years beyond K-means,” Pattern Recognit. Lett., vol. 31, no. 8, pp. 651–666, 2010.

[3] A. K. Jain, M. N. Murty, and P. J. Flynn, “Data Clustering – A Review,” ACM Computing Surveys. vol. 31, no. 3, pp. 265-323, Sept. 1999.

[4] P-N Tan, M. Steinbachm, and V. Kumar, “Introduction to Data Mining,” PEARSON Addison Wesley, 2006.

[5] C.Zahn. “Graph-theoretical methods for detecting and describing gestalt clusters”. IEEE Transactions on Computers, C-20:68-86, 1971.

[6] F.Preparata and M.Shamos. “Computational Geometry: An Introduction”. Springer-Verlag, Newyr, NY ,USA, 1985

[7] T.Asano, B.Bhattacharya, M.Keil and F.Yao. “Clustering Algorithms based on minimum and maximum spanning trees”. In Proceedings of the 4th Annual Symposium on Computational Geometry,Pages 252-257, 1988

[8] Mahlagha Sedghi and George K. Atia, "A Multi-Criteria Approach for Fast and Robust Representative Selection from Manifolds", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 34, NO. 7, JULY 2022.

[9] Gaurav Mishra and Sraban Kumar Mohanty, "RDMN: A Relative Density Measure Based on MST Neighborhood for Clustering Multi-Scale Datasets", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 34, NO. 1, JANUARY 2022.

[10] Y.Xu, V.Olman and D.Xu. “Minimum spanning trees for gene expression data clustering”. Genome Informatics,12:24-33, 2001.

[11] Q. Cheng, X. Lu, Z. Liu, J. Huang, and G. Cheng, “Spatial clustering with density-ordered tree,” Physica A: Statistical Mechanics Appl., vol. 460, pp. 188–200, 2016

[12] J. Rojas-Thomas, M. Santos, and M. Mora, “New internal index for clustering validation based on graphs,” Expert Syst. Appl., vol. 86, pp. 334–349, 2017.

[13] Eggels, A., & Crommelin, D. (2019). Quantifying data dependencies with Rényi mutual information and minimum spanning trees. Entropy, 21 (2). DOI 10.3390/e21020100

[14] Sieranoja, S., & Fränti, P. (2019). Fast and general density peaks clustering. Pattern Recognition Letters, 128 , 551–558. DOI 10.1016/j.patrec.2019 .10.019

[15] Gagolewski, M. (2021). genieclust: Fast and robust hierarchical clustering. SoftwareX, 15 , 100722. DOI 10.1016/j.softx.2021.100722

[16] Ma, Y., Lin, H., Wang, Y., Huang, H., He, X. (2021). A multi-stage hierarchical clustering algorithm based on centroid of tree and cut edge constraint. Information Sciences, 557 , 194–219. DOI 10.1016/j.ins.2020.12.016

[17] Amit Kumar Das, Basabi Chakraborty, Saptarsi Goswami, Amlan Chakrabarti, "A fuzzy set based approach for effective feature selection", Fuzzy Sets and Systems, Volume 449, 2022.

[18] Campagner, A., Ciucci, D., Denœux, T. (2023). A general framework for evaluating and comparing soft clusterings. Information Sciences, 623 , 70–93. DOI 10.1016/j.ins.2022.11.114

[19] F. Pasi et al., "Clustering datasets," 2015. [Online]. Available: http://cs.uef.fi/sipu/datasets/

[20] C. C. Yeh and M. S. Yang, "Evaluation measures for cluster ensembles based on a fuzzy generalized Rand i0ndex," Appl. Soft Comput., 2017

[21] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in Proc. Int. Conf. Knowl. Discovery Data Mining, 1996, pp. 226–231.

[22] J. Xie, Z.-Y. Xiong, Y.-F. Zhang, Y. Feng, and J. Ma, "Density corebased clustering algorithm with dynamic scanning radius," Knowl.-Based Syst., vol. 142, pp. 58–70, 2018.

[23] M.Halkidi, Y. Batistakis, andM.Vazirgiannis, "On clustering validation techniques," J. Intell. Inf. Syst., vol. 17, no. 2–3, pp. 107–145, 2001.

[24] Kshirsagar, P. R., Reddy, D. H., Dhingra, M., Dhabliya, D., & Gupta, A. (2023). A scalable platform to collect, store, visualize and analyze big data in real- time. Paper presented at the Proceedings of 2023 3rd International Conference on Innovative Practices in Technology and Management, ICIPTM 2023, doi:10.1109/ICIPTM57143.2023.10118183 Retrieved from www.scopus.com

[25] Kshirsagar, P. R., Reddy, D. H., Dhingra, M., Dhabliya, D., & Gupta, A. (2022). Detection of liver disease using machine learning approach. Paper presented at the Proceedings of 5th International Conference on Contemporary Computing and Informatics, IC3I 2022, 1824-1829. doi:10.1109/IC3I56241.2022.10073425 Retrieved from www.scopus.com

[26] Kumar, A., Dhabliya, D., Agarwal, P., Aneja, N., Dadheech, P., Jamal, S. S., & Antwi, O. A. (2022). Cyber-internet security framework to conquer energy-related attacks on the internet of things with machine learning techniques. Computational Intelligence and Neuroscience, 2022. doi:10.1155/2022/8803586