Available online at https://sanad.iau.ir/journal/ait
Publisher : Ya.C., Islamic Azad University
Artificial Intelligence Tools in Software and Data Engineering

Enhancing Reliability Prediction of Power Transformers in the Petrochemical Industry using Deep Learning Models

Komail Pourkhaghan[1], Mehran Emadi[2]*
[1]Master Student, Department of Electrical Engineering, Mo.C., Islamic Azad University, Isfahan, Iran.
Email: komail.pourkhaghan@iau.ir
[2]Department of Electrical Engineering, Mo.C., Islamic Azad University, Isfahan, Iran.*
Email: mehranemadi49@iau.ac.ir (Corresponding author) ***orcid.org/0000-0001-6976-4250***

Article type: Research Article (Article)

Article history: Received , , Received in revised , , Accepted , Published online ,

***Abstract*— This research uses advanced deep learning and data mining methods to analyze and improve the reliability of power transformers in the petrochemical industry. The main goal is to identify critical equipment and optimize maintenance and repair strategies by using fuzzy clustering algorithms and fuzzy distance functions. For this purpose, transformer performance data including repair costs, mean time between failures (MTBF), downtime, and criticality score were collected and, after preprocessing, were entered into unsupervised machine learning algorithms. In the analysis process, the data were categorized into five separate clusters, which increased the accuracy in identifying critical equipment. Simulation results showed that the optimal number of clusters for reliability analysis is five clusters. The model evaluation using Mean Squared Distance (MSD) and Objective Function (OBJ) metrics showed that the fuzzy clustering algorithm performed better than the traditional methods. Specifically, the OBJ value decreased by about 12% and MSD by about 10%, which is due to the use of fuzzy distance functions to more accurately represent the relationships between data. Finally, this research provides practical suggestions for improving equipment management, including focusing on preventive maintenance of critical equipment, optimizing repair schedules for old and expensive equipment, and maintaining stable performance through regular maintenance. This approach can lead to cost reduction, increased productivity, and improved reliability of power systems in the petrochemical industry.**

***Index Terms*— Reliability Analysis, Power Transformers, Deep Learning, Fuzzy Clustering, Fuzzy Distance Functions, Petrochemical Industry.**

I. INTRODUCTION

IN TODAY'S complex and competitive world, ensuring reliability in power supply system's especially in critical industries such as petrochemicals is of paramount importance. Power transformers, as key components of electrical networks, play a vital role in the transmission and regulation of electrical energy. Failures in these devices can lead to production shutdowns, significant financial losses, and even threats to personnel and environmental safety [1–3]. Therefore, precise reliability analysis and timely fault prediction are considered fundamental priorities in asset management within this industry [4]. Traditional methods for evaluating transformer reliability mainly rely on offline inspections, diagnostic testing, and expert judgment. However, these approaches face serious limitations in scalability, accuracy, and real-time applicability [4]. With the growing volume of industrial data and the advancement of artificial intelligence particularly deep learning new data-driven approaches have emerged for identifying critical equipment, modeling failure patterns, and optimizing maintenance plans [5, 6]. Reliability Centered

Cite this article: Komail Pourkhaghan, Mehran Emadi,. (2025). Enhancing Reliability Prediction of Power Transformers in the Petrochemical Industry using Deep Learning Models . ***Journal of Artificial Intelligence Tools in Software and Data Engineering (AITSDE),*** (1), pages.

Maintenance (RCM) is recognized as a structured framework aimed at improving asset performance by aligning maintenance activities with actual reliability requirements [7]. Successful implementation of RCM depends on the accurate identification of critical equipment, allowing for optimal and targeted repair planning. However, conventional RCM approaches often based on questionnaires and historical records from specific plants lack generalizability to other operational conditions [8, 9]. In the petrochemical industry, power transformers play a crucial role in supplying electricity for production processes, and their failure can result in considerable financial and operational losses. The technical and economic complexity of replacing such equipment underscores the importance of preventive maintenance and effective management. Given the increasing number and variety of transformers in this sector, traditional repair planning methods based on fixed schedules and past experiences no longer provide sufficient efficiency. Accordingly, adopting intelligent, data-driven methods such as machine learning and deep learning has become essential for maintenance optimization and risk reduction [10, 11]. To overcome these limitations, this study proposes an innovative approach using unsupervised deep learning and fuzzy clustering algorithms with fuzzified distance functions to improve the reliability analysis of power transformers in the petrochemical industry [12]. In this research, operational data including repair costs, Mean Time Between Failures (MTBF), downtime, and equipment criticality scores were collected and preprocessed before being fed into clustering algorithms [13–14]. The proposed approach employs a fuzzified similarity function capable of representing hidden and nonlinear relationships between data with higher accuracy. By clustering transformers into five distinct groups based on performance and criticality criteria, the method enables more precise and targeted maintenance decisions. Evaluation results show a 12% improvement in the OBJ metric and a 10% reduction in MSD, indicating superior reliability modeling accuracy compared to traditional methods [15, 16]. Given the growing need for intelligent solutions in critical industries such as petrochemicals to enhance the reliability of electrical equipment, this research introduces a novel framework based on unsupervised deep learning, taking a significant step toward advancing traditional maintenance and fault analysis methods for power transformers. The innovations of this study can be summarized in three main aspects:

Integration of fuzzy clustering algorithms with fuzzified distance functions: Unlike conventional approaches that use classical distances (e.g., Euclidean), this study utilizes fuzzified distance functions to more accurately analyze complex relationships among operational features in uncertain environments. This improves clustering precision and critical equipment identification.

Practical application of deep learning in the petrochemical industry for critical equipment identification: While most prior research has focused on power generation or transportation sectors, this study applies fuzzy clustering based on machine learning to real-world petrochemical data, evaluating its effectiveness under complex industrial conditions.

Development of a generalizable framework for intelligent decision-making in preventive maintenance: The proposed model is not only applicable for diagnosing power transformer conditions in petrochemical plants but can also be extended to other critical assets and similar industries. This flexibility makes it a practical tool for implementing predictive maintenance across various industrial environments.

The remainder of this paper is organized as follows. Section 2, presents a review of related work. Section 3, provides the necessary definitions and details of the proposed method. Section 4, evaluates the proposed method, and Section 5, concludes the paper with a comprehensive summary.

II. RELATED WORK

Limited efforts have been made to identify the most critical electrical equipment in various industries, particularly in the petrochemical sector. Typically, identifying critical equipment is specifically carried out to enhance system reliability. Reference [17], was the first to systematically apply RCM to power distribution systems. However, identifying key equipment in the process has often been reduced to a simple sensitivity analysis, in which some of the essential criteria involved in the problem are overlooked. Consequently, RCM has been criticized as being overly broad and generalized. In [18], an empirical framework was presented to reduce corrective maintenance in hydraulic systems within the context of Lean 4.0. This study used an action research methodology in nine stages to analyze and optimize the maintenance process. The results indicated that implementing Lean 4.0 had a direct impact on maintenance performance and could improve understanding and actions in maintenance management through a qualitative approach. The framework effectively linked Lean 4.0 with practical research to enhance maintenance in manufacturing, including petrochemical industries. In [19], a machine learning-based method was introduced for identifying critical equipment in the aluminum sheet industry. Data collection included failure history and maintenance records of selected components. An exponential distribution was used to analyze failure trends, reliability, and mean time between failures (MTBF), alongside maintenance costs. The proposed method improved maintenance cost reduction by 3.6% while achieving accurate identification of failure-prone equipment. Reference [20], proposed an adaptive ARIMA-based model for predicting failures in electrical components, reducing crises in Industry 3.0 by up to 84%. The model, implemented on a high-pressure hydraulic sand molding machine, analyzed key parameters such as oil contamination levels. It increased the mean time between failures by up to 800%, with strong performance in minimizing unexpected downtimes. In [21], an RCM-based asset management model was proposed for developing maintenance strategies for power transformers. A Markov model was used to estimate optimal reliability and cost solutions. Sensitivity analysis was performed on transition rates, and real service data were used to estimate transformer lifespan. In [22], a fuzzy Analytic Hierarchy Process (AHP) was implemented for asset management in power distribution systems. This study emphasized reliability as a central factor in asset management, focusing on demand optimization through RCM. The role of fuzzy systems in asset management was thoroughly analyzed. In [23], a topological analysis of electrical transmission networks was conducted using metrics such as global efficiency and meshedness to assess vulnerability. Here, vulnerability refers to changes in power flow or losses, and the goal is to identify critical components to improve stability and protect the system against natural or human induced threats. In [24], the dynamics of persistent topological features were analyzed under node and edge removal scenarios using various deletion strategies. A weighted graph with a predefined threshold was considered, where nodes with lower weights were removed. The VR method was used to extract topological features from the transformed graph, and the proposed approach was evaluated in the Texas transmission distribution system. In [25], an improved method for transformer fault diagnosis based on a residual BP neural network was proposed. This approach deepened the BP network using stacked residual modules and integrated SVM for filtering high-accuracy feature vectors. Gas features were analyzed using deep mapping and multi-layer fusion, with the highest-weighted feature vector selected for final fault classification. The proposed model achieved a 92% fault diagnosis accuracy, outperforming traditional BP networks. A review of previous studies reveals that most existing methods for transformer reliability analysis, based on machine learning and systematic modeling, have improved failure prediction and maintenance efficiency. However, many of these methods depend heavily on precise data and predefined assumptions, which may not be available in real-world conditions. Additionally, models such as AHP and topological vulnerability analysis, despite offering practical benefits, face challenges like long processing times and structural limitations. Thus, hybrid approaches based on deep learning and data mining capable of managing complex datasets are necessary to overcome these constraints.

III. PROPOSED METHOD

A-Definitions Required to Solve the Problem

Definition 1. The symbol $Sim(x_i.x_j)$ is used to represent the degree of clustering similarity. The desired data set is defined as $x = \{x_1. \dots . x_n\}$, where the similarity between each pair of data x_i and x_j is defined using equation 1.

$$Sim_{x_i.x_j \in X}\left(x_i.x_j\right) = \lambda d_s\left(x_i.x_j\right) + (1-\lambda)\left|\cos\left(x_i.x_j\right)\right| \quad (1)$$

Where $d_s\left(x_i.x_j\right)$, is the fuzzy Euclidean distance defined by equation (2) [41].

$$d_s\left(x_i.x_j\right)_{x_i.x_j \in X} = f\left(\frac{d(x_i.x_j) - \mathrm{Min}\{d(x_i.x_j)\}}{\mathrm{Max}\{d(x_i.x_j)\} - \mathrm{Min}\,\{d(x_i.x_j)\}}\right) \quad (2)$$

And the symbol $d(x_i.x_j)_{x_i.x_j \in X}$ in equation 2 is the same as the Euclidean distance $\sqrt{\sum_{k=1}^{m}(x_i - x_j)^2}$. The coefficient λ is the weight factor and its value varies between zero and one and takes a different value in vectors. But in experiments we consider it equal to 0.5. The symbol $\cos(x_i.x_j)_{x_i.x_j \in X}$ is the cosine of the angle between two vectors and can be calculated through equation 3, and its value varies between [-1.1][27].

$$\cos(x_i.x_j)_{x_i.x_j \in X} = \frac{\sum_{i=1}^{m}(x_i x_j)}{\sqrt{\sum_{i=1}^{m} x_i^2}\sqrt{\sum_{j=1}^{m} x_j^2}} \quad (3)$$

Usually, for clustering analysis, the distance criterion or similar criteria is used to measure the similarity of vectors. Also, to calculate the similarity degree of the vector set x, the lower triangular matrix $SimMatrix[n]$can be used, because for the similarity degree of both vectors, $Sim(x_i.x_j)$ and $Sim\left(x_j.x_i\right)$ are equivalent to each other.

Definition 2. The similarity neighborhood of the vector x_i in the vector set x is represented by $SimNeighbor(x_i.\alpha)$. For each data x_i, we assume that x_i is the cluster center, and we compare the clustering similarity degree for each data with the data x_i. If this value is greater than the threshold value α, the data in question will be the similarity neighborhood of x_i. This function is defined as equation 4 [42].

$$\mathrm{SimNeighbor}(x_i.\alpha)_{x_i \in X} = \{x \mid \alpha \le \mathrm{sim}(x_i.x)_{x_i.x \in X} \le 10 \le \alpha \le 1. X = \{x_1. \dots . x_n\}\} \quad (4)$$

Definition 3. The similarity density criterion of vector x_i, which is calculated using equation 5, is denoted as $Density(x_i)$.

$$\mathrm{Density}(x_i) = \frac{\sum_{j=1}^{\left|P_{\mathrm{neighbor}(x_i)}\right|} \mathrm{Sim}(x_i.P^{j}_{\mathrm{neighbor}(x_i)})}{\left|P_{\mathrm{neighbor}(x_i)}\right|}. X = \{x_1. \dots . x_n\}. \quad (5)$$

The symbol $P^{j}_{neighbor(x_i)} \in SimNeighbor(x_i.\alpha)$ indicates that these vectors satisfy the threshold α in the neighborhood of similarity x_i. The value $\left|P_{neighbor(x_i)}\right|$ indicates the number of data that exist in the neighborhood of data x_i.

Definition 4. Clustering similarity indicates the average similarity degree of vectors belonging to a group. Suppose there are two data objects x_i and x_j, the clustering similarity degree has a value between zero and one $0 \le Sim(x_i.x_j) \le 1$ .the clustering similarity of the cluster C_p is expressed as equation 6. The closer the data in a cluster are to each other, the larger their clustering similarity value will be [43].

$$\mathrm{Hom}(C_p) = \frac{1}{|C_p|^2} \sum_{i=1}^{|C_p|} \sum_{j=1}^{|C_p|} \mathrm{Sim}(x_i . x_j)_{x_i . x_j \in X} \tag{6}$$

Definition 5. For two given clusters C_p and C_q, the clustering Homogeneity measure represents the average similarity of points belonging to two different clusters. The clustering heterogeneity is calculated using equation 7.

$$\mathrm{Het}(C_p . C_q) = \frac{1}{|C_p||C_q|} \sum_{i=1}^{|C_p|} \sum_{j=1}^{|C_q|} \mathrm{Sim}(x_i . x_j)_{x_i . x_j \in X} \tag{7}$$

Definition 6. The main goal of clustering is the optimal division of a data set. In this process, data that have certain and defined similarities are placed in the same group (cluster) as much as possible. Conversely, data that have little similarity with each other are assigned to different groups. This approach helps to identify hidden patterns and structures in the data. The evaluation coefficient of clustering results obj or target is defined as relationship 8 [43].

$$\mathrm{obj} = \mathrm{Compactness}(C) \times \mathrm{Seperation}(C) \tag{8}$$

In equation 8, the values of Compactness and Separation are calculated using equations 9 and 10, respectively.

$$\mathrm{Compactness}(C) = \sum_{i=1}^{k} \frac{\mathrm{Hom}(C_i)}{k} \tag{9}$$

$$\mathrm{Separation}(C) = 1 - \sum_{i=1}^{k} \sum_{j=1}^{k} \frac{\mathrm{Het}(C_i . C_j)}{k} \tag{10}$$

Evaluating the quality of clustering is not a difficult task. The quality of clustering can be evaluated using the density within each cluster and the amount of difference between clusters. The Compactness variable indicates the similarity of the members of a cluster, which is calculated using equation 9. The separation criterion, which indicates the amount of dissimilarity of data belonging to different groups, is calculated using equation 10. Our goal is to maximize the density within a cluster and minimize the similarity between data belonging to different clusters.

Definition 7. As stated earlier, the similarity criterion in this study is determined based on the fuzzy Euclidean distance between points. This criterion states that the smaller the Euclidean distance between two points, the greater their similarity. Accordingly, if the Euclidean distance of points within a cluster is smaller, the value of the density criterion increases. On the other hand, if the distance between points belonging to different clusters decreases, according to equation 11, the value of the separation criterion increases. In this case, the MSD criterion also increases, which indicates a better quality of clustering [43].

$$\mathrm{MSD} = \sum_{K} \frac{\sum_{i=1}^{m^{(k)}} \| d_i - d^{(k)} \|^2}{m^{(k)}} \tag{11}$$

The proposed method includes a step-by-step framework for identifying critical equipment in transformer maintenance management, which is based on fuzzy clustering and fuzzy distance functions. First, operational and expert data are prepared after preprocessing including cleaning, normalization, and removal of invalid values. Then, a matrix of key transformer characteristics such as failures, energy losses, repair costs, and lifespan is formed. Next, the fuzzy clustering algorithm is implemented using fuzzy distance functions to classify equipment and is iteratively optimized until the objective function is reached. Finally, by analyzing the clustering results, critical equipment is identified and prioritized, and optimal maintenance programs are proposed for them. This approach will increase the accuracy of reliability analysis and reduce operating costs. Figure 1, shows the outline of the proposed method.

III.I. Preprocessing

In this research, data preprocessing includes identifying and replacing missing values using the KNN method, removing noise and human errors, and normalizing the data to the range [0,1]. These steps were necessary to improve the quality of real data collected from petrochemical industry transformers. Finally, the cleaned data were prepared as a feature matrix for input into the fuzzy clustering algorithm [28]. In this method, if x_{min} and x_{max} are the minimum and maximum data values recorded, respectively. In this case, in the range y_{min} and y_{max}, which are the minimum and maximum values desired in normalization. The following relationship expresses normalization in the data.

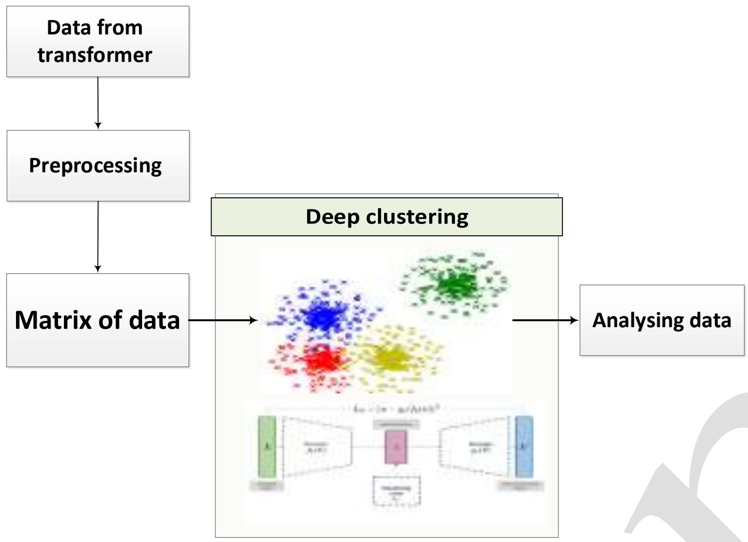

Fig 1. Steps of the proposed method.

$$y_i = y_{\min} + (y_{\max} - y_{\min}).\frac{(x_i - x_{\min})}{x_{\max} - x_{\min}} \tag{12}$$

Since in this research, $y_{\min} = 0$, $y_{\max} = 1$ and equation 12, is reconstructed as equation 13.

$$y_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{13}$$

Since this transformation is linear, the distribution of the data does not change.

III.II. Improving clustering with FCM algorithm with fuzzy distance functions

In this research, to improve clustering accuracy, the Fuzzy Clustering Means algorithm (FCM) along with fuzzy distance functions is used. This approach reduces the value of the objective function in each iteration by gradually updating the cluster centers and membership degrees and prevents convergence to local solutions. Using fuzzy distance instead of traditional criteria increases the ability to model complex nonlinear relationships in real data and leads to more accurate results.

$$E = \sum_{j=1}^{k} \sum_{i=1}^{n} f\left(\left\| x_i^{(j)} - c_j \right\|^2\right) \tag{14}$$

In Equation 14, f is defined as the distance fuzzification function. To improve the performance of the FCM algorithm, the distance fuzzification function is used as a key component in this study. This function allows modeling complex and nonlinear relationships in real data and improves the accuracy and quality of clustering. Using this approach not only reduces the impact of random selection of initial cluster centers, but also increases the ability of the algorithm to identify hidden features and complex data structures. To improve the performance of the FCM algorithm and reduce the instability caused by random selection of initial cluster centers, this study uses fuzzy distance functions and a white box approach. The number of clusters is determined by analyzing the data structure and consulting with experts to prevent incorrect selection. Also, an evaluation coefficient is designed to measure the quality of clustering, and the clustering process is continued in an exploratory and gradual manner until the desired stability is achieved.

III.III. FCM Clustering Steps with Fuzzy Functions

The main idea of the FCM algorithm with fuzzy distance functions is to calculate the similarity matrix Sim Matrix[n] for the m-dimensional vector X. The similarity neighborhood matrices and similarity density are also calculated with respect to the Sim Matrix. Initially, the IDs of the initial center candidates are placed in an array called X'. These points are the same as the initial center points. In the array X', the point x_i with the highest similarity density in the determined similarity neighborhood is selected as one of the K initial cluster centers and recorded in init C. By selecting each point x_i, all points that satisfy the Sim Neighbor(x_i.α) constraint are removed from the array X'. In a similar way, the number of K initial cluster centers is determined and recorded in init C. The selection of K initial cluster centers is iteratively as follows:

Algorithm input: Array X containing n data elements with m dimensions, array X' to indicate the ID of candidate points, similar neighborhood threshold α, similarity weighting factors λ and the set of initial clustering centers.
Algorithm output: Number of K clusters that form the set C, such that the relationships $C_i \cap C_j = \emptyset,\ i \neq j.0 \leq i.j \leq K$ hold between the clusters. Table 1, presents the process in a structured and step-by-step manner to help in easier understanding and use in algorithm development.

TABLE I. Pseudo code of FCM algorithm with fuzzy distance functions.

Step	Description of Operations
Input	Array X: contains n data elements with m dimensions Array X′: stores the IDs of candidate points α: Similarity neighborhood threshold λ: Similarity weighting factor init C: set of initial clustering centers
Output	Set C: contains K clusters such that $C_i \cap C_j = \emptyset$ there is no overlap between clusters such that the relations $C_i \cap C_j = \emptyset$, $i \neq j.0 \leq i.j \leq K$ Compute the Sim Neighbor (x_i.α) matrix for the similarity between all data points x_i in X.
Calculate similarity matrix	Calculate the Sim Neighbor (x_i.α) matrix for the similarity between all data points x_i in X
Calculate adjacency matrix	For each point x_i: A. Identify points that are in the neighborhood of x_i and satisfy the constraint Sim Neighbor (x_i.α). B. Calculate the similarity density of each point.
Initialize array X′	Fill the array X′ with the IDs of the candidate points.
Select initial cluster centers	In an iterative loop: a. Select the point x_i with the highest similarity density in X.′ b. Store x_i in init C c. Remove all points that satisfy the Sim Neighbor (x_i .α) constraint from X,′
Repeat step 4	Continue the process until the number K of initial cluster centers in init C is determined.
Final output	The output contains the initial cluster centers of init C and the set of clusters C.

III.IV. Proposed Deep Clustering

Given an input dataset $X = \{x_i;\ x_i \in R^D\}_{i=1}^{N}$, the learning goal of the embedded autoencoder is to construct a better encoder f(0; ω₁), which makes the latent representations $z_i \in R^{D'}(D' < D)$ more suitable for clustering. In this way, the input encoder unit maps $\overline{x_i} \in R^D$ to the latent representation $z_i \in R^{D'}(D' < D)$ in the low-dimensional space with a nonlinear mapping f(0; ω_1), as shown in equation (16). The decoder unit reconstructs the input t$\overline{x_i} \in R^D$ from its latent representation z_i with the nonlinear mapping $g_{\omega 2}: z_i \rightarrow \overline{x_i}$, as shown in equation 2. Here, ω_1 and ω_2 are the learnable parameters (weight and bias) for the encoder and decoder units, respectively.

$$z_i = f_{w1}(x_i) \tag{15}$$

$$\overline{x_i} = g_{w2}(z_i) \tag{16}$$

Figure 2, shows the architecture of an embedded deep auto encoder for deep clustering, which consists of two nested encoder and decoder sections. In this architecture, each section contains an internal auto encoder that performs multiple encoding-decoding steps to produce more accurate latent representations of the data. The internal encoder in the encoder section extracts high-level features of the input data before final encoding of the data. In contrast, the embedded decoder in the final reconstruction step performs decoding-encoding operations to improve the reconstruction of the hidden layers, thereby providing more robust latent representations for clustering. Deep clustering is performed on the extracted latent representations $Z = \left\{ z_i;\ z_i \in R^{\acute{D}} \right\}_{i=1}^{N}$ to cluster them into K clusters. The soft assignment used for this purpose is defined as the similarity between the embedded observations z_i and the cluster center μ_j. To achieve this goal, the distribution matrix $Q = [q_{ij}]_{N*K}$ is used. The matrix Q consists of q_{ij} elements, which are measured by the t distribution [12] as follows.

$$q_{ij} = \frac{(1+(\|z_i-\mu_j\|^2)^{-1}}{\sum_{j=1}^{k}(1+(\|z_i-\mu_j\|^2)^{-1}} \tag{17}$$

Where q, is the probability of assigning observation i to cluster j. The auxiliary distribution matrix $P = [p_{ij}]_{N*K}$is introduced to improve cluster purity and emphasize observations assigned with high confidence.
It is given by the formula:

$$p_{ij} = \frac{q_{ij}{}^2 / \psi_j}{\sum_{j=1}^{k} q_{ij}{}^2 / \psi_j} \tag{18}$$

Where $\psi_j = \sum q_{ij}$, denotes the soft cluster frequencies. Clustering is performed by the substitution between the soft label Q and the auxiliary target distribution P. The clustering loss is defined as the KL (Kullback– Leibler) divergence [13] between the two distributions: the soft label Q and the target distribution P[3]. This clustering loss is calculated by minimizing the following equation:

$$L_{KLD} = KL(P \| Q) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}} \tag{19}$$

Where L_{KLD}, is the clustering loss. Then, the predicted label of the observation $z_i \in Z$ corresponding to $x_i \in X$ is assigned to class j, which satisfies the following equation:

$$l_i = argmax_j q_{ij} \tag{20}$$

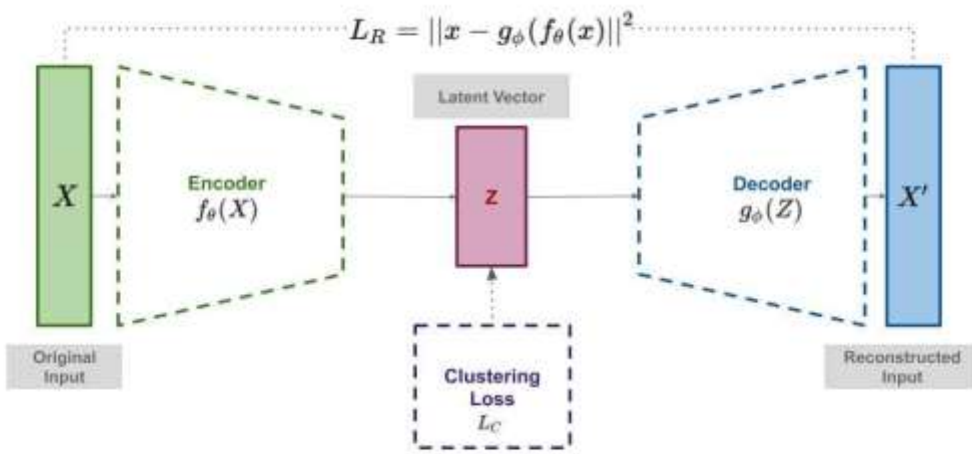

Fig 2. Generalized architecture of embedded deep autoencoder for deep clustering.

IV.RESULTS

In this study, a system equipped with an Intel Core i7 processor (Q720 @ 2.60GHz), Windows 10 operating system, and MATLAB 2021b software were used to simulate and execute the proposed algorithm. This hardware and software configuration enabled simultaneous processing of a large volume of data and execution of complex fuzzy clustering and machine learning algorithms. The use of these tools has led to accurate, fast, and reliable analyses on real data from the petrochemical industry, which helps to improve the identification of critical equipment, increase system stability, and reduce operational risks.

A. Research Database

The research database contains a comprehensive set of operational and environmental characteristics of power transformers in the petrochemical industry. The data includes input and output voltage, load current, power consumption, failure type and time, repair cost and time, and environmental conditions such as temperature and humidity. Information on the type of maintenance (preventive or corrective) is also included in the collection. This data is collected with the aim of reliability analysis, failure prediction, and maintenance process optimization. The precise structure of this database allows the use of machine learning algorithms to identify critical equipment and develop operational risk reduction solutions. Table 2, shows the data specifications.

B. Evaluation Criteria

Since this study presents a method for improving clustering with distance fuzzification functions, it seems that in the first stage, the degree of clustering improvement should be evaluated and then the proposed method is analyzed in risk assessment and the resulting clusters. Therefore, the method has been implemented on the collected data and clustered. The OBJ criteria [29] and the mean sum of squares of the distance between the points and the centers of the clusters MSD [30] have been used for clustering.

C. Determining the Number of Clusters

The Elbow Method plot shows the amount of Distortion, or sum of squared errors within a cluster, for different numbers of clusters. Skewness is a measure of how close the data are to their cluster centers. As the number of clusters increases, the amount of skewness decreases because each cluster becomes smaller and the data gets closer to its centers. In this plot, a significant reduction in skewness is observed at first, but after a certain point, the reduction in skewness continues to be gradual and imperceptible. This point is known as the Elbow point and indicates the optimal number of clusters. Based on the plot presented, the Elbow point can be considered as the optimal number of clusters for data analysis. Figure 3, assumes the optimal number of clusters to be 5. However, the analysis will be performed for the entire number of clusters.

TABLE II. Data specifications.

Feature	Range/Values	Data type	Explanation
Transformer ID	T1, T2, T3,...,T500	Text	Each transformer ID
Input Voltage (KV)	110, 132, 150, 220	Number	Transformer input voltage (KV)
Output Voltage (KV)	33, 66, 75, 132	Number	Transformer output voltage (KV)
Load Current (A)	300 to 800	Number	Load current applied to the transformer (Amperes)
Power (MVA)	40 to 100	Number	Power converted by the transformer (Megavolt-Amperes)
Failure Time hours	1000 to 2500	Number	Transformer failure time (in hours)
Failure Type	"Insulation failure", "Winding failure", "Cooling failure", "Other"	Text	Type of transformer failure (insulation, winding, cooling, etc.)
Repair Time Hours	4 to 10	Number	Transformer repair time (in hours)
Repair Cost	4000 to 8000	Number	Transformer repair cost (currency)
Ambient Temperature C	30 to 50	Number	The ambient temperature in which the transformer is located (degrees Celsius)
Humidity_%	50 to 70	Number	Humidity of the environment where the transformer is located (percentage)
Maintenance Type	"Preventive Corrective"	Text	Type of maintenance performed on the transformer

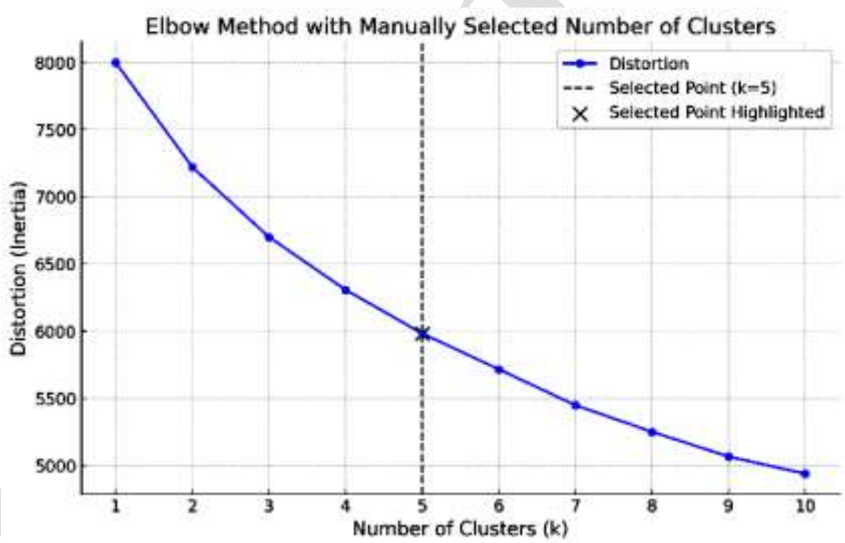

Fig 3. Determining the appropriate number of clusters.

D. Clustering Evaluation with OBJ

Figure 4, shows the clustering results for different numbers of clusters in the proposed method. As can be seen from this figure, by increasing the number of clusters towards the 5 clusters predicted by the expert, the obj criterion also improved significantly. When the number of clusters increases to 10, the obj criterion decreases compared to 5 clusters. This shows that the number of 5 clusters is sufficient and appropriate for assessing the existing risks. By examining the graph in figure 4, it is clear that in all cases, the obj criterion for the proposed method is lower than that of the traditional method. This decrease in the obj value indicates an improvement in the clustering quality using the improved algorithm. Especially in clustering with a smaller number (such as 4 or 5 clusters), the difference between the two methods is more noticeable, which indicates that the proposed method performs much better in this situation. Also, with the increase in the number of clusters, both methods show a gradual increase in the obj criterion value. This increase is natural, because the increase in the number of clusters leads to a more complex distribution of the data. However, even in more complex conditions (such as 9 or 10 clusters), the proposed method still performs better than the traditional method. In conclusion, this graph proves that the use of fuzzy distance functions in the FCM algorithm reduces errors and improves the clustering quality, and the proposed method is more efficient than the traditional algorithm.

- **Comparison of methods based on the sum of the mean squares of the distances to the cluster center**

The lower the sum of the mean squares of the Euclidean distances of the points from the centers of the clusters, the better the clustering result; because it indicates a high density of points within the obtained clusters. Figure 5 shows the distance criterion in the proposed method. By increasing the clusters to 5 clusters, this criterion has shown a decrease, while the distance has shown an increase with increasing the number of clusters. The distribution of the available data as well as the risks defined in the data set can be a major reason for justifying this issue.

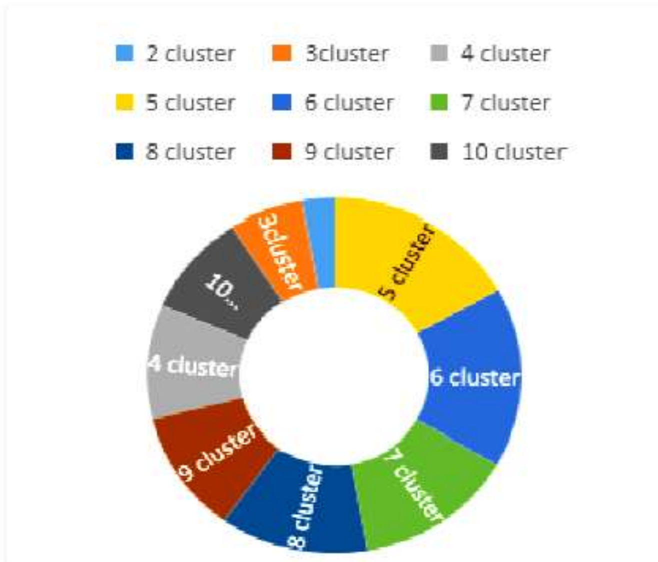

Fig 4. Comparison chart of obj criterion in different number of clusters in the proposed method.

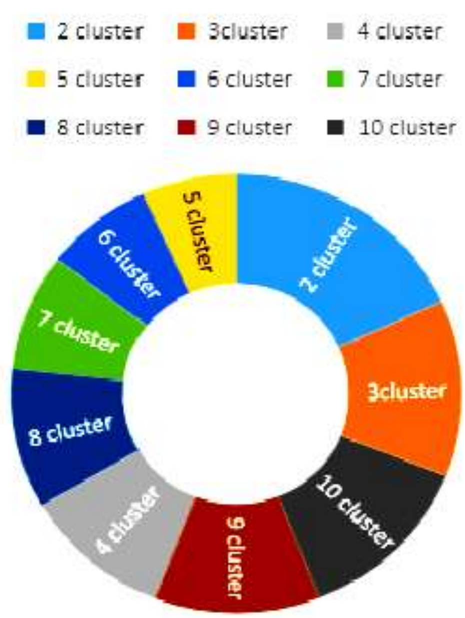

Fig 5. Comparison chart of distance criterion in number of clusters in the proposed method.

In figure 5, the values obtained from the calculation of this criterion after running both clustering methods on the test dataset are compared with each other. Unlike the previous criterion, these values differ more in the number of clusters. So that in the lower number of clusters, the MSD value in the FCM method is very different from the improved method. Therefore, this method is not suitable for clusters with less than 5 clusters. By increasing the number of clusters from 6 clusters to 10 clusters, the difference in the MSD value in the methods decreases, while the MSD value has increased compared to 5 clusters. However, in each execution of the methods with any number of clusters, the improved FCM algorithm has provided the lowest MSD value and, as a result, the best clustering result. Of course, the smaller the distance between the points within each cluster and the center of that cluster, the more similar (or close) the points within each cluster are to each other, which will lead to an increase in density within each cluster.

E. Comparison of MSD and OBJ

The obj criterion represents the sum of errors or the distance of the data from the cluster centers. In the graph related to this criterion, in all the numbers of clusters (from 2 to 10), the proposed method has performed better than the traditional FCM method and has shown a lower value for the obj criterion. This indicates that the data in the proposed method are closer to their cluster centers and the data distribution in the clusters is more optimal. The decrease in the obj criterion is very evident in the number of fewer clusters (such as 4 and 5 clusters), which indicates the high effectiveness of the proposed method in simpler conditions. The distance criterion also examines the proximity of the data to the cluster centers, but from a different angle. The graph related to this criterion shows that in all the numbers of clusters, the proposed method has a lower distance criterion than the traditional method. In this criterion, too, the difference between the two methods is greater in the number of fewer clusters (such as 4 and 5 clusters). This shows that the proposed method not only increased the clustering accuracy, but also managed to bring the data closer to the cluster centers.

F. Analysis of Results

According to the information in Table 3, which includes the analysis of various characteristics for the five transformer clusters in the petrochemical industry, a more detailed technical analysis of each cluster can be provided. This analysis is based on various criteria such as repair cost, time between failures, downtime, failure probability, and criticality score. The technical analysis of each cluster is presented below:

- **Cluster 0 (transformers with stable performance and low maintenance requirements)**

This cluster includes equipment that is in a relatively stable state in terms of repair cost ($5234.23) and time between failures (245.87 hours). The repair cost in this cluster is moderate compared to other clusters, and the downtime (12.54 hours) also indicates a relative impact on the overall system performance. The probability of failure of these transformers is also moderate (0.075), meaning that failures do not occur frequently. The criticality score (7.25) indicates the moderate importance of these equipment in maintaining the overall system performance. This cluster requires periodic inspection, but overall, its condition is stable and cannot be given a high priority for repair and maintenance.

- **Cluster 1 (Transformers with high repair costs and need for preventive maintenance)**

This cluster includes equipment with the highest repair costs ($6,745.12) and relatively long service life (14.7 years). Due to the high probability of failure (0.085) and the longest downtime (14.76 hours), these equipment require more detailed and preventive maintenance programs. These transformers are in a more critical condition and should be prioritized in maintenance programs. The high criticality score (8.14) indicates that the failure of these equipment can have a large impact on system performance. This cluster requires regular actions and more detailed checks to reduce the risk of failure.

- **Cluster 2 (Most Stable Transformers)**

Cluster 2, includes equipment with the lowest repair costs ($4,523.67) and the longest time between failures (275.89 hours). These transformers have stable performance with low failure probability (0.065) and low downtime (11.32 hours). This cluster can be considered as an example of equipment with optimal performance that does not require special attention. The equipment of this cluster is technically very efficient and does not normally require special maintenance to continue operation. In terms of cost, these transformers are the least expensive equipment to repair.

- **Cluster 3 (Most Critical Transformers)**

This cluster contains the equipment with the highest criticality score (8.76) and the highest probability of failure (0.092). The cost of repairs for these equipment is high ($7,123.45) and their downtime is longer (15.87 hours), indicating the special importance of these equipment in maintaining system performance. These transformers require immediate and careful attention in maintenance and repair programs to prevent possible failures. Failure of these transformers can have the greatest impact on production and efficiency and therefore should be given high priority for maintenance.

- **Cluster 4 (Medium Performance Transformers)**

Cluster 4, contains equipment with repair costs ($4,932.87) and time between failures (260.78 hours) in the medium range. These transformers require attention compared to other clusters but are not as critical as cluster 3. The criticality score (7.56) and probability of failure (0.072) of these equipment indicate that they are moderately impactful and should be monitored and inspected regularly. The downtime (13.45 hours) also indicates that this equipment has a moderate impact on overall performance.

TABLE III. Review of critical clusters.

Criterion	Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4

Repair Cost (mean)	5234.23	6745.12	4523.67	7123.4	4932.8
Repair Cost (std)	1503.45	1345.78	1234.56	1402.6	1298.4
MTBF Hours (mean)	245.87	198.34	275.89	210.56	260.7
MTBF Hours (std)	38.25	45.12	40.34	50.78	36.4
Dow time Hours (mean)	12.54	14.76	11.32	15.87	13.4
Downtime Hours (std)	3.87	4.22	3.54	4.56	3.76
Failure Probability (mean)	0.075	0.085	0.065	0.092	0.07
Failure Probability (std)	0.012	0.014	0.011	0.016	0.013
Inspection Interval (mean)	6.34	5.67	7.12	5.45	6.78
Inspection Interval (std)	1.25	1.34	1.12	1.56	1.23
Operating Hours (mean)	4920.34	5234.12	4765.89	5402.3	4876.4
Operating Hours (std)	567.89	678.45	498.12	589.67	534.5
Age Years (mean)	12.4	14.7	11.3	15.2	12.9
Age Years (std)	4.2	3.9	4.5	4.1	3.8
Criticality Score (mean)	7.25	8.14	6.87	8.76	7.56
Criticality Score (std)	1.45	1.78	1.34	1.92	1.67
Failure Count (mean)	4.87	5.12	4.34	5.67	4.76
Failure Count (std)	1.23	1.56	1.08	1.34	1.15
Repair Cost (mean)	5234.23	6745.12	4523.67	7123.4	4932.8

V. CONCLUSION

In this study, a novel method for improving reliability analysis in power transformers in the petrochemical industry based on a combination of deep learning and fuzzy clustering was presented. Using fuzzy distance functions, clustering accuracy was increased and more accurate identification of critical equipment in maintenance processes was provided. Real operational data from equipment in the petrochemical industry was used and the clustering results were analyzed using evaluation indices such as MSD and OBJ. Evaluations showed that using 5 clusters provides the best balance between analysis accuracy and data segmentation significance and optimizes decision-making in transformer maintenance. This finding has direct application in developing targeted maintenance strategies, reducing operating costs, and increasing the stability of power systems. Finally, based on the results obtained, suggestions for further improvement were presented; Including the use of deeper learning models to increase the accuracy of failure prediction, the application of new performance evaluation indicators, and the development of intelligent maintenance frameworks based on environmental and operational data. This approach can play a key role in improving productivity and reducing operational risks in the petrochemical industry.

REFERENCES

[1] R. O. Okeke, A. I. Ibokette, O. M. Ijiga, L. A. Enyejo, G. I. Ebiega, and O. M. Olumubo, “THE RELIABILITY ASSESSMENT OF POWER TRANSFORMERS,” Engineering Science & Technology Journal, vol. 5, no. 4, pp. 1149-1172, 2024.

[2] G. lal Rajora, M. A. Sanz-Bobi, and C. M. Domingo, “Application of Machine Learning Methods for Asset Management on Power Distribution Networks,” Emerging Science Journal, vol. 6, no. 4, pp. 905-920, 2022.

[3] J. R. Geisbush, Using Reliability Centered Maintenance (RCM) Analyses toDevelop Large Diameter Water Pipeline Maintenance Strategies, Arizona State University, 2024.

[4] K. Zadiran, and M. Shcherbakov, “New Method of Degradation Process Identification for Reliability-Centered Maintenance of Energy Equipment,” Energies, vol. 1 ,۶no. 2, pp. 575, 2023.

[5] I. H. Afefy, A. Mohib, A. El-kamash, and M. Mahmoud, “A New Framework of Reliability Centered Maintenance,” Jordan Journal of Mechanical & Industrial Engineering, vol. 13, no. 3, 2019.

[6] Seyed Abolghasemi, S.S., Emadi, M. and Karimi, M., 2024. Accuracy improvement of breast tumor detection based on dimension reduction in the spatial and edge features and edge structure in the image. *Majlesi* Journal of Electrical Engineering, *18*(1), pp.33-44.

[7] J. Geisbush, and S. T. Ariaratnam, “Reliability centered maintenance (RCM :(literature review of current industry state of practice,” Journal of Quality in Maintenance Engineering, vol. 29, no. 2, pp. 313-337, 2023.

[8] B. Wang, B. Fang, Y. Wang, H. Liu, and Y. Liu, “Power system transient stability assessment based on big data and the core vector machine,” IEEE Transactions on Smart Grid, vol. 7, no. 5, pp. 2561-2570, 2016.

[9] H. Ghasemi, E. S. Farahani, M. Fotuhi-Firuzabad, P. Dehghanian, A. Ghasemi, and F. Wang, “Equipment Failure Rate in Electric Power Distribution Networks: An Overview of Concepts, Estimation, and Modeling Methods,” Engineering Failure Analysis, pp. 107034, 2023.

[10] L. Duchesne, E. Karangelos, and L. Wehenkel, “Recent developments in machine learning for energy systems reliability management,” Proceedings of the IEEE, vol. 108, no. 9, pp. 1656-1676, 2020.

[11] Yang, Hsin-Jung, Mahsa Khosravi, Benjamin Walt, Girish Krishnan, and Soumik Sarkar. "Zero-shot Sim-to-Real Transfer for Reinforcement Learning-based Visual Servoing of Soft Continuum Arms." arXiv preprint arXiv:2504.16916 (2025).

[12] Al Hamadani, Rihab Abdullah Jaber, et al. "Improvement of Network Traffic Prediction in Beyond 5G Network using Sparse Decomposition and BiLSTM Neural Network." *Qubahan Academic Journal* 5.2 (2025): 156-176.

[13] A. AIROBOMAN, “Reliability Assessment of Power System Network: A Detailed Review ”,International Journal of Engineering Technologies IJET, vol. 7, no. 4, pp. 90-103.

[14] S. H. Hoseinie, U. Kumar, and B. Ghodrati, Reliability centered maintenance (RCM) for automated mining machinery: Luleå tekniska universitet, 2016.

[15] Karimi, Mohsen, Zahra Karimi, Mahsa Khosravi, Zeinab Delaram, Mostafa Habibi Dehsheikhim, Somayeh Arab Najafabadi, Mohammadreza Alizadeh Aliabadi, and Nakisa Tavakoli. "Feature selection methods in big medical databases: a comprehensive survey." *International Journal of Theoretical & Applied Computational Intelligence* (2025): 181-209."
[16] S. K. Sahu, M. Roy, S. Dutta, D. Ghosh, and D. K. Mohanta, “Machine learning based adaptive fault diagnosis considering hosting capacity amendment in active distribution network,” Electric Power Systems Research, vol. 216, pp. 109025, 2023.
[17] L. Bertling, “Reliability-centred maintenance for electric power distribution systems,” Elektrotekniska system, 2002.
[18] N. M. d. M. Torre, and A. Bonamigo, “Action research of lean 4.0 application to the maintenance of hydraulic systems in steel industry,” Journal of Quality in Maintenance Engineering, 2024.
[19] S. Bani, and H. Nwosu, “Assessing the Sustainability of an Aluminum Roofing Sheet Corrugating Machine by Reliability Centered Maintenance and Cost Reduction Technique”.
[20] T. R. Mohan, J. P. Roselyn, R. A. Uthra, D. Devaraj, and K. Umachandran, “Intelligent machine learning based total productive maintenance approach for achieving zero downtime in industrialmachinery,” Computers & Industrial Engineering, vol. 157, pp. 107267, 2021.
[21] A. Koksal, and A. Ozdemir, “Improved transformer maintenance plan for reliability centred asset management of power transmission system,” IET Generation, Transmission & Distribution, vol. 10, no. 8, pp. 1976-1983, 2016.
[22] K. Ayu, and A. Yunusa-Kaltungo, “A holistic framework for supporting maintenance and asset management life cycle decisions for power systems,” Energies, vol. 13, no. 8, pp. 1937, 2020.
[23] D. Bose, C. K .Chanda, and A. Chakrabarti, “Vulnerability assessment of a power transmission network employing complex network theory in a resilience framework,” Microsystem Technologies, vol. 26, no. 8, pp. 2443-2451, 2020.
[24] B. Li, D. Ofori-Boateng, Y. R. Gel, and J. Zhang, “A hybrid approach for transmission grid resilience assessment using reliability metrics and power system local network topology,” Sustainable and Resilient Infrastructure, vol. 6, no. 1-2, pp. 26-41, 2021.
[25] Y. Jin, H. Wu, J. Zheng, J. Zhang, and Z. Liu, “Power transformer fault diagnosis based on improved BP neural network,” Electronics, vol. 12, no. 16, pp. 3526, 2023.
[26] L. Rutkowski, M. Jaworski, and P. Duda, "Basic Concepts of Data Stream Mining," Stream Data Mining: Algorithms andTheir Probabilistic Properties, pp. 13-33: Springer, 2020.
[27] R. Gao, J. Yin, R. Liu, Y. Liu, J. Li, L. Dong, W. Ma, L. Zhang, P. Zhang, and Z. Tian, “A novel particle size distribution correction method based on image processing and deep learning for coal quality analysis using NIRS-XRF,” Talanta, vol. 285, pp. 127427, 2025.
[28] M. R. Malarvizhi, and A. S. Thanamani, “K-nearest neighbor in missing data imputation,” International Journal of Engineering Research and Development, vol. 5, no. 1, pp. 5-7,2012.
[29] M. K. Gupta, and P. Chandra, "An Empirical Evaluation of K-Means Clustering Algorithm Using Different Distance/Similarity Metrics," Proceedings of ICETIT 2019, pp. 884-892: Springer, 2020.
[30] K. R. Devi, “Evaluation of Partitional and Hierarchical Clustering Techniques,” 2019.