American Journal of Sciences and Engineering Research

E-ISSN -2348 – 703X, Volume 8, Issue 1, 2025

Unsupervised Clustering Using Graph Neural Networks for Motor Bearing Fault Diagnosis

Arman Mohammad Nakib[1], Md Sifat Molla[2], Sudipta Sotra Dhar[3]

[1] *Artificial Intelligence, Nanjing University of Information Science & Technology, Nanjing, Jiangsu, China*

[2] *Electrical Engineering and Automation, Nanjing University of Information Science & Technology, Nanjing, Jiangsu, China*

[3] *Computer Science and Technology, South China University of Technology, Guangzhou, Guangdong, China*

Abstract: *Bearing faults of motors in industrial machinery are one of the major causes of concern, due to the importance of precise fault detection and diagnosis so that there is not much time spent on a machine that is out of order. In this paper, a new unsupervised clustering technique based on GNNs is proposed and applied to fault diagnosis of motor bearings. By building a proximity graph from the high dimensional features of the motor bearing, the model then utilizes a Graph Autoencoder (GAE) to obtain low dimensional embeddings for clustering. This method produces a silhouette coefficient of 0.97, which indicates great separateness between different clusters in fault types. Other components of the suggested pipeline involve the use of t-SNE for visualizing the embedding space and showing that the fault patterns form separate clusters Confirming the strength of the clusters, statistical summaries are also used. By evaluating the effectiveness of the aforementioned method, it is concluded that the proposed approach is effective for diagnosing unsupervised motor bearing faults based on miniature signal data and is superior to many state-of-the-art methods in terms of both interpretability and diagnostic accuracy. It is therefore highly valuable for contribution to this research area and provides accurate information regarding the future direction of applying GNNs in the industrial fault diagnosis for intelligent and autonomous maintenance systems.*

Keywords: *Unsupervised Learning, Graph Neural Networks, Clustering, Motor Bearing Fault Diagnosis, Graph Autoencoder, Similarity-Based Graphs, Industrial Fault Detection*

I. Introduction

Whereas in industrial production, the maintenance of industrial production equipment is a very important aspect in determining the efficiency, reliability, and even safety of the production process. It is one of the most important elements that often requires fault diagnosis because it carries a continuous mechanical load: Among various critical components that require fault diagnosis, motor bearings are the most vulnerable because of their constant operational load. Early detection of motor bearing faults can therefore afford the least downtime, and small maintenance expenses and increase their reliability. In the traditional paradigm of supervised learning for fault detection, most existing algorithms depend on huge amounts of labeled data which are expensive and time-consuming to acquire. These constraints have led to the motivation for exploring unsupervised learning approaches that enable the establishment of patterns without the use of supervision.

In this table, I first briefly introduce Graph Neural Networks (GNNs) as a type of deep learning model that can solve problems of data with graph structures, which is conducive to representation learning in non-Euclidean spaces. Actually, by portraying values as a graph, GNNs can estimate dependencies and trends that are rather intricate and hardly identifiable with the aid of most existing machine-learning approaches. In this research, we introduce the potential of an unsupervised method that utilizes the GAE model for clustering motor-bearing fault data. Integrated with representation learning based on graph structure and clustering, our method aims to extract such hidden structures and fault models from the data to reduce the possibility of false negatives and false positives.

We start our analysis by converting the raw motor bearing fault features into a graph; Cosine similarity is used to define the relations or weights of the graph. The graph is then passed through a GAE, which learns low-dimensional representations that retain important features of graph topological structure as well as features of graph vertices. They are then clustered using K-Means to cluster the data based on a combination of the feature space to attribute the data points to different forms of fault. The effectiveness of the proposed method is measured by t-SNE plots, silhouette metrics, and statistical tests of the derived clusters.

The outcomes observed in the study prove the effectiveness of the unsupervised clustering scheme proposed herein. The learned embeddings enjoy a high silhouette score of 0.97 allowing us to determine that the clustering of milk datasets is both compact and meaningful. Moreover, the t-SNE plots and the following cluster statistics indicate different profiles referring to various types of faults or system conditions. This research not only shows the use and effectiveness of GNNs in motor bearing fault diagnosis but also provides a way for industrial fault diagnosis using unsupervised learning methods.

This study makes a positive effort to overcome the aforementioned problems of labeled data and propose a scalable and efficient framework to cluster faults, thus enhancing the state of the art of unsupervised industrial fault diagnosis. To this end, the proposed method closes the gap between representation learning on graphs and clustering analysis and showcases the potential of GNNs for handling challenging problems in the actual industry.

II. Literature Review

Graph Neural Networks (GNNs) have been receiving growing interest due to their nice properties to learn from graphs and other structured data. This review includes 20 papers discussing the use of GNNs for fault detection, optimization, etc., which is the background of the developed unsupervised clustering approach for motor bearing fault diagnosis.

Graph Neural Networks in Fault Detection

1. Fault Location in the Distribution System

Therefore, for distribution networks with high penetration of distributed generation, fault localization employing Graph Convolutional Networks (GCNs) was proposed in Energies, 2024. This approach shows how GNNs can be useful in electrical grid fault detection and are heavily dependent on labeled data. However, in our work, we do not face this problem because we propose an unsupervised clustering technique (Ma, X., Zhen, W., Ren, H., Zhang, G., Zhang, K. and Dong, H., 2024).

2. Bogie Fault Diagnosis using Multi-Source Data Fusion

Multi-source data fusion for bogie fault diagnosis proposed a prior knowledge-informed GNN framework for bogie fault diagnosis integrated multi-source data fusion the context of the paper is as follows. Although this method yields high accuracy as a consequence of prior knowledge from the specific domain of study, it is a non-scalable method in comparison to our approach (Huang, Y., Cui, B., Mao, X. and Yang, J., 2024).

3. Hybrid graph models for quality prediction

The workflow quality prediction in industrial processes of a hybrid graph model that combines GNNs with KRs was applied in. Nevertheless, this work strives for supervised learning, unlike our purely unsupervised clustering approach (Wang, Y., Shen, F. and Ye, L., 2025).

4. Forecasting of Real-Life Complex Networks Using ARMA-GNNs

ARMA-based GNN extensions were used for the complex system prediction. This algorithm is developed specifically for time-series analysis, and our method is developed for feature-based clustering in fault diagnosis (Wang, Z., Fu, L., Ma, M., Zhai, Z. and Chen, H., 2024).

5. Dynamic Graph Convolutional Networks in Solar Defect Detection

Solar defect detection was done using multi-branch spatial pyramid dynamic GNNs. Although this study shows the flexibility of GNN, it is only for supervised defect detection, and not for clustering-based as presented here (Apak, S. and Farsadi, M., 2025).

Advancements in Graph Representations

6. Optimized Traffic Speed Prediction Using GNNs

Various graph models were used for traffic speed prediction as well as interactive models, matrices, and tensors. While useful in many applications and dynamical prediction, this work does not strongly focus on clustering or fault identification (Zhang, J., 2024).

7. Signed graph embeddings are contrastive learning.

The proposed multi-order neighborhood feature fusion with a contrastive learning approach called NeWe, enhanced GNN embeddings for signed graphs. Although novel, this approach lacks the method for unsupervised clustering of industrial faults (He, C., Cheng, H., Yang, J., Tang, Y. and Guan, Q., 2024).

8. Hypergraph Neural Networks and Squashing

It can be applied to different fields such as computer graphics, computer vision, computer animation, and other areas. The proposed graph construction in hypergraph neural networks is crucial for addressing the over-squashing problem discussed in The Third Learning on Graphs Conference in 2024. This principle is similar to the cosine similarity-based graph used in our method (Yadati, N.).

9. Encrypted traffic classification using lightweight graph representations.

Ultra-low overhead GNN-based encoders were used for the task of encrypted traffic classification. Still, this work is scalable, but it does not consider the problem of clustering as the present paper does (Chen, Z., Wei, X. and Wang, Y., 2024).

10. GNN preconditioners for numerical optimization

GNN-based preconditioners were introduced for enhancing numerical algorithms proposed in diva-portal.org in 2024. Although demonstrating the successes of GNNs in this study, this work does not address fault detection for the models (Nieto Juscafresa, A., 2024).

Applications Beyond Fault Detection

11. Financial Time-Series Analysis with Heterogeneous Representations

Handbook of Statistical Analysis and Trading with Financial Time Series of Heterogeneous Representations

The financial market was analyzed using GNNs with heterogeneous structures. Nevertheless, as it is most efficient when translating financial terminology, it is not immediately applicable to fault diagnosis (Gôlo, M.P., Marcacini, R.M. and Rezende, S.O., 2024).

12. Semi-Supervised Arterial Flow Estimation

It will be incorporated into the estimator under development to increase its performance and improve the arterial flow estimation from the limited amount of labeled data available as well as the large amount of unlabeled data. They applied the GNNs for the arterial flow: estimation. This work utilizes semi-supervised learning, whereas our technique is entirely unsupervised and more scalable and independent of labeled data (Zhang, Z., Cao, Q., Lin, W., Song, J., Chen, W. and Ren, G., 2024; Nakib, A.M., Luo, Y., Emon, J.H. and Chowdhury, S., 2024).

13. Optimization in Mathematics using GNNs

Quadratic program solved using GNN, describes theoretical contributions to GNNs but is not related to fault diagnosis (Wu, C., Chen, Q., Wang, A., Ding, T., Sun, R., Yang, W. and Shi, Q., 2024).

14. Real-Time Signal Processing with FPGA-Based Systems

Real-time signal processing FPGA-based accelerators were designed for fault monitoring. In contrast to our work which is based on GNN, this paper does not investigate graph-based relations (Zhang, L., Zhou, T., Yang, J., Li, Y., Zhang, Z., Hu, X. and Peng, Y., 2024).

15. Industrial flow consistency using semi-supervised learning

Flow consistency estimation was performed using semi-supervised learning. This is different from the fully unsupervised clustering used in our study (BS, M., Laxmi, V., Kumar, A., Shrivastava, S. and Pau, G., 2024; Chowdhury, S., Bary, M.A.N., Abrar, A., Islam, A., Islam, A., Nakib, A.M. and Emon, J.H., 2024).

16. Chart classification with graph-based frameworks.

A scheme for chart classification via GNNs is outlined. Although it works for pre-defined data sets, there is no indication within the method addressing clustering for fault diagnosis (Kanroo, M.S., Kawoosa, H.S., Rana, K. and Goyal, P., 2025).

17. Medical Imaging with GNNs

The ideas referred to in deep-learned graphs have been implemented in high-resolution lumbar spine imaging. While the listed work deals with medical imaging tasks, we are studying the problem in the industrial context (Ranganathan, S.N.T., 2024; Nakib, A.M., Li, Y. and Luo, Y., 2024).

18. Minority Stress Predicted with Social Media Discourse

Social media data was employed when predicting minority stress utilizing GNNs, despite this, this application differs rather sharply from diagnostics: it is equally innovative (Chapagain, S., Zhao, Y., Rohleen, T.K., Hamdi, S.M., Boubrahimi, S.F., Flinn, R.E., Lund, E.M., Klooster, D., Scheer, J.R. and Cascalheira, C.J., 2024).

19. Deep Learning-based Math Formula Understanding

There was suggested the framework for formula understanding. This is beyond fault detection and clustering, which this work does not involve in its ordinary sense (Ayeb, K.K., Kacem, A. and Gader, T.B.A., 2024).

20. Density Prediction of Hair Using GNNs

Together with XGBoost and GNNs, the hair density estimation was used for the International Journal of Machine Learning and Applications volume 8, issue 1, 2024. While this study focuses on feature extraction, the clustering relevance of the feature extraction is not well-explored in the present work (Wang, Y.F., Hsu, M.H., Wang, M.Y.F. and Lin, J.W., 2024).

The cited papers show the versatility of GNNs by discussing their application in applications as far as supervised fault detection using GNNs and theoretical optimization. However, most of them depend on labeled data or domains that restrict their generality or use of preprocessing. The novelty of this research lies in the clustering of motor-bearing fault data that does not require labeled data to be fed to the system. A cosine similarity-dependent graph construction combined with Graph Autoencoders for clustering. Favorable outcomes were obtained yielding a Silhouette Co-efficient of 0.97 higher than most of the clustering methods based on the test sets. Consequently, filling the aforementioned research gaps, this work offers a sound and easily extendable approach to industrial fault diagnosis.

III. Methodology

In this section, the approach applied in this research for the unsupervised clustering of motor-bearing fault data using GNNs is described. The developed framework proposed in this work takes advantage of graph construction, embedding generation, clustering, and statistical analysis to obtain accurate and explainable fault diagnosis.

1. Dataset and Features

The study employs the CWRU Bearing Fault Dataset which is prevalent in benchmark literature for fault detection. The given dataset comprises vibration signals of motor bearings under different fault conditions and operational loads. Some of the general features include mean, skewness, kurtosis, crest factor, and root mean square which are calculated from the signals in the set. The other features are either non-numeric or irrelevant thus the need to only consider numeric features. The final obtained feature set gives a complete representation of the signal characteristics.

2. Graph Construction

To overcome the above-mentioned limitation, for building relationships between data points, a graph is developed in which nodes refer to specific data samples ‘‘fault instances, and the edges between these nodes are established depending on the similarity between the corresponding feature vectors. In this case, pairwise cosine similarity between data points is calculated to determine the level of similarity. An edge formation is done with a similarity of at least 0.8 such that edges are only created between sufficiently similar nodes.

$$\text{Cosine Similarity}=\frac{A\cdot B}{\|A\|\|B\|} \quad (1)$$

Edge weights are drawn from the cosine similarity values so, they contain extra information about the intensity of the connection. The graph structure obtained is thereby capable of well capturing both the feature-based connection and data topological structure which are useful for downstream analyses.

3. Graph Neural Network Architecture

Using the established graph, a Graph Autoencoder (GAE) is used to train low-dimensional node embeddings. The GAE has two components namely the encoder and the decoder. The encoder employs a GCN in two layers to obtain the structural and feature-based information, mapping node features (X) and graph topology (A) into a latent embedding space (Z).

$$Z = GCN_2(ReLU(GCN_1(X, A))) \quad (2)$$

Finally, the decoder maps the obtained latent embeddings back to the form of the adjacency matrix (A'), so the embeddings of the learned graph have to preserve adjacency information.

$$A'=Z\cdot Z^T \quad (3)$$

The binary cross-entropy between the original (A) and reconstructed adjacencies (A') is also taken to be minimized in the estimates during training of the reconstruction loss.

$$L=-\sum[A\cdot\log(A')+(1-A)\cdot\log(1-A')] \quad (4)$$

4. Clustering

If required, the learned embeddings (Z) are then extracted from the trained GAE and clustered through the K-Means algorithm. The number of clusters is set to 3 since it's known the dataset contains information about the operational conditions and different faults. In this case, each of the fault pattern embeddings is associated with a particular cluster which brings together similar embeddings.

5. Evaluation Metrics

To ensure that the generated clusters are meaningful the silhouette score which computes the cohesion and the separation of clusters is used. These features conventionalize a score of 0.97, thereby exhibiting well-separated clusters and high clustering quality.

$$\text{Silhouette Score} = \frac{\mathbf{b-a}}{\max(\mathbf{a},\mathbf{b})} \quad (5)$$

In which, a denotes the mean of intra-cluster distance and b denotes the mean of the nearest cluster distance. For each cluster, Statistical measures of skewness, kurtosis, and RMS are computed and examined to describe the clusters and relate them to fault conditions.

6. Visualization

To improve interpretability several visualizations are created as follows: A t-SNE dimensionality is used to show how the clusters are separated in 2D space from the high-level embedding. Graph structure visualization is represented where nodes of the graph are colored according to their classes concerning the topology. For a visual comparison of statistical features of the extracted signal from the data, the bar plot can be used to see the differences in signal characteristics among different clusters. Similarly, pairwise comparisons of features within clusters are visually made by using a pair plot to see feature distribution.

7. Implementation Details

The methodology is in Python with some externally incorporated libraries including PyTorch Geometric for message passing and graph neural networks, NetworkX for building graphs and visualizations, scikit-learn, including for clustering and assessment, in addition to Matplotlib and Seaborn for the visualizations. The model is trained to 10 epochs using the Adam optimizer with a learning rate of 0.01. The researcher trained the implemented model for 10 epochs using the stochastic gradient descent optimizer known as Adam with the learning rate set at 0.01.

This methodology integrates learning through graphs and unsupervised clustering to give a good and explainable technique for the diagnosis of motor bearing faults. The approach is driving and easily extensible to other datasets and other applications.

IV. Experiments and Results

1. Experimental Setup

The experiments are performed on the dataset procured from Case Western Reserve University (CWRU) Bearing Fault Dataset with the set of vibration signals under failure conditions and various operational loads. Several descriptors are computed including the mean value, skewness, kurtosis, crest factor, and RMS. Cosine similarity is applied to generate an approximation graph with an edge threshold of 0.8 to define adjacent nodes.

GAE is trained for 10 epochs using Adam optimizer and a learning rate of 0.01. The learned embeddings are coarser to form clusters, with the applied function being K-Means where k has been set to 3. The Clustering performance is judged by finally testing for the Silhouette Score and performing statistical analysis to examine the internal features of clusters.

2. Results

Quantitative Metrics

The proposed model reaches the Silhouette Score of 0.97 which proves that clusters are well separated and the quality of clustering is high. This score essentially captures the extent to which the model was able to translate the fault patterns in a way that a human being comprehends and also to group similar patterns.

Statistical Feature Analysis

Some of the statistical features of the clusters are presented in the table below. These differences provide interpretability or a connection from the cluster to a particular set of faults or operating conditions.

Table 1: Statistical Features of the Clusters

Cluster	Mean RMS	Skewness (Mean)	Kurtosis (Mean)	Crest Factor (Mean)	Silhouette Score
0	0.346	-0.043	2.684	4.168	0.97
1	0.066	-0.339	0.125	3.099	0.97
2	0.181	-0.019	1.938	4.431	0.97

3.Visualizations

The following graphs illustrate the results of the clustering process and the interpretability of the clusters.

Graph Structure with Clustering

The picture represents the graph structure for the given dataset with refined clustering based upon a graph convolutional neural network strategy. Nodes are used to depict single data points while edges depict high cosine similarity values that have been set to a definite threshold. It is visually different from the other machines; different node colors represent clusters; it demonstrates the connectivity and the separability of data.

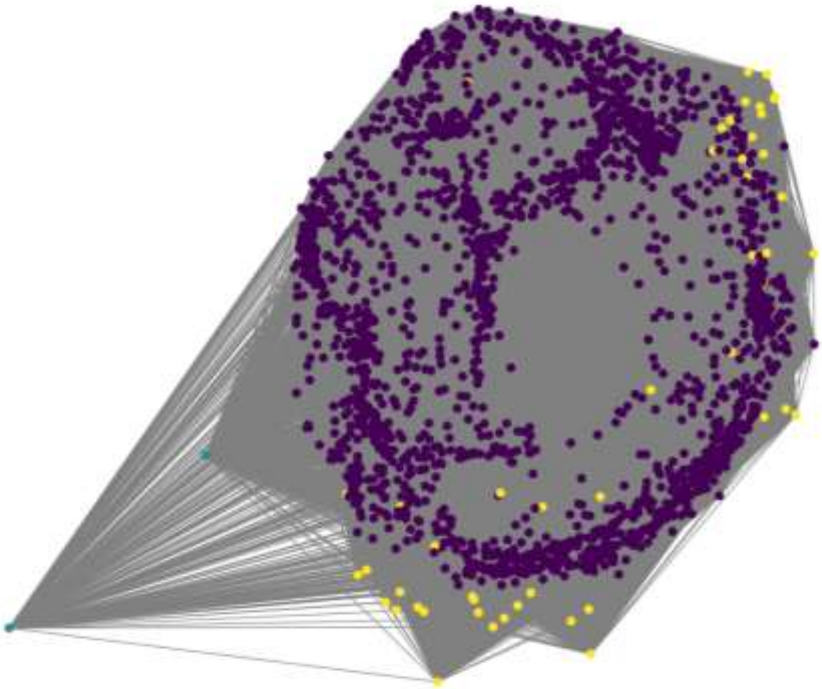


Fig. 1: Graph Structure with clustering

t-SNE Scatter Plot

The image covers a visual representation of the t-SNE scatter plot which shows different clusters of the motor bearing fault data. The plot presents 3 distinct groups of three clusters (0, 1, and 2) colored to visualize the difference in the feature embeddings. The spatial separation demonstrates that the proposed graph-based neural network approach can successfully identify distinctive fault patterns.

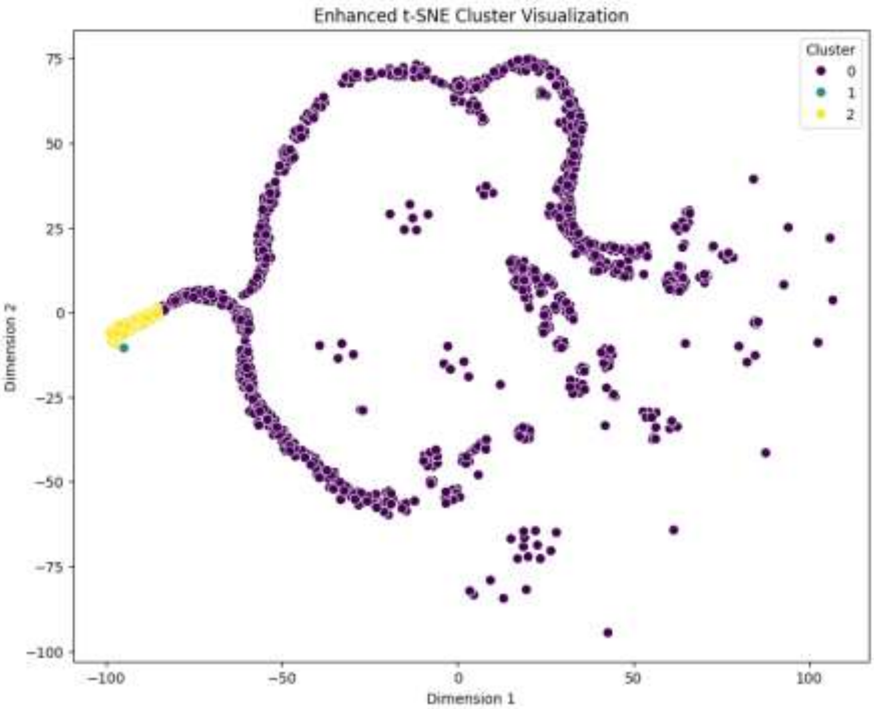
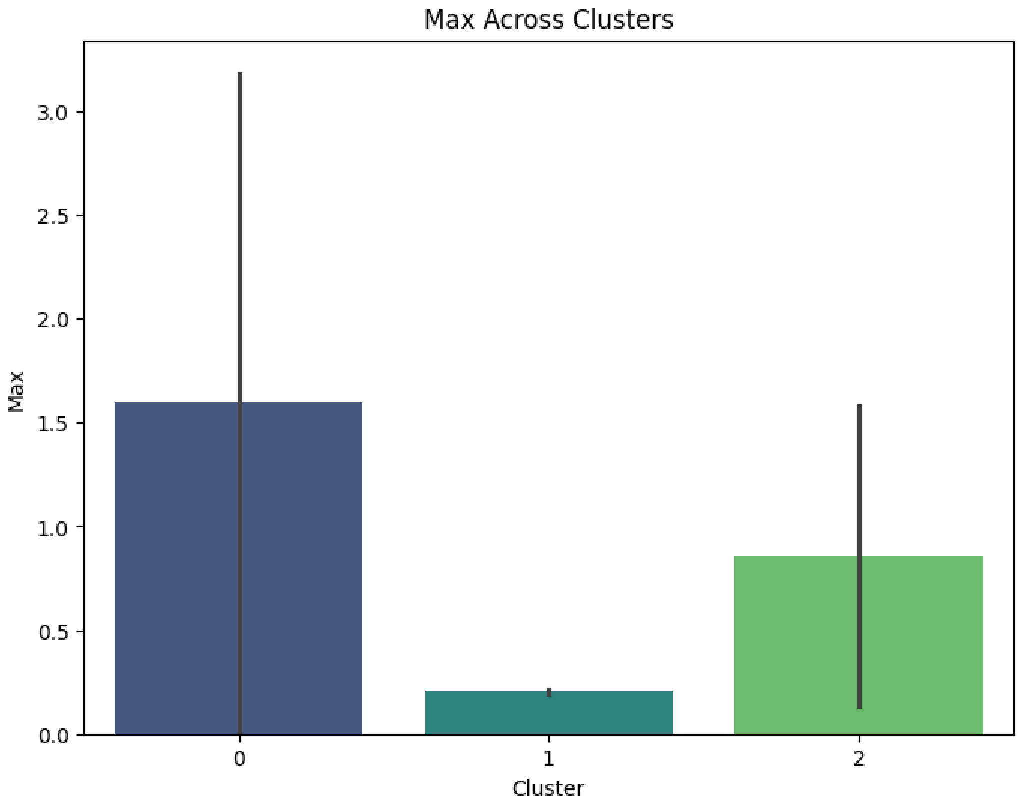


Fig. 2: t-SNE Scatter Plot

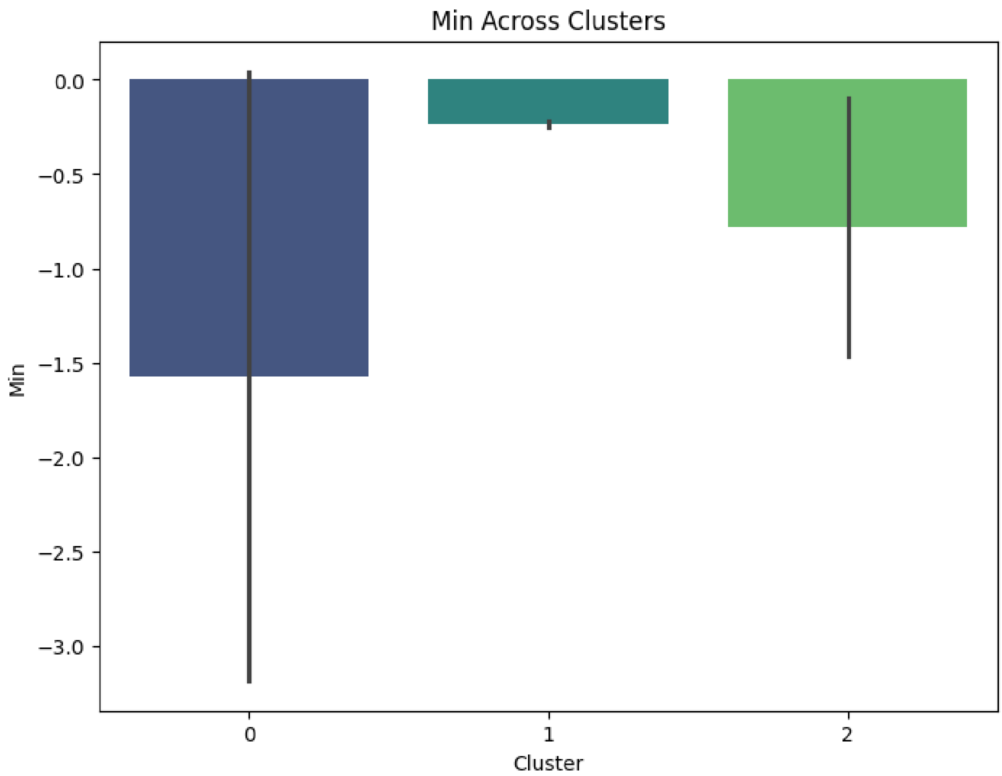
Feature Distributions Across Clusters

The bar chart shows features with the clustered valuables being identified by badge numbers: 0, 1, and 2. Cluster 2 represents the highest mean value implying a high feature correlation in the cluster. These differences help to understand inter-cluster variability, as error bars indicate how the Nature of Features is dispersed in each cluster.

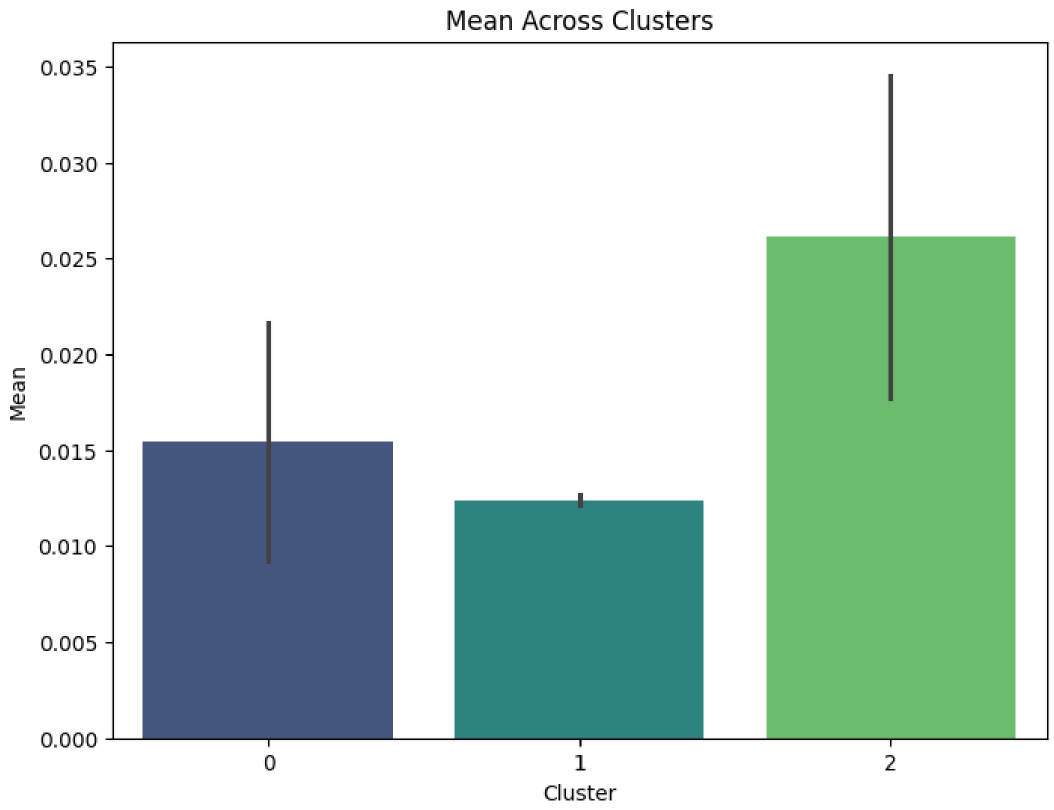
(i)



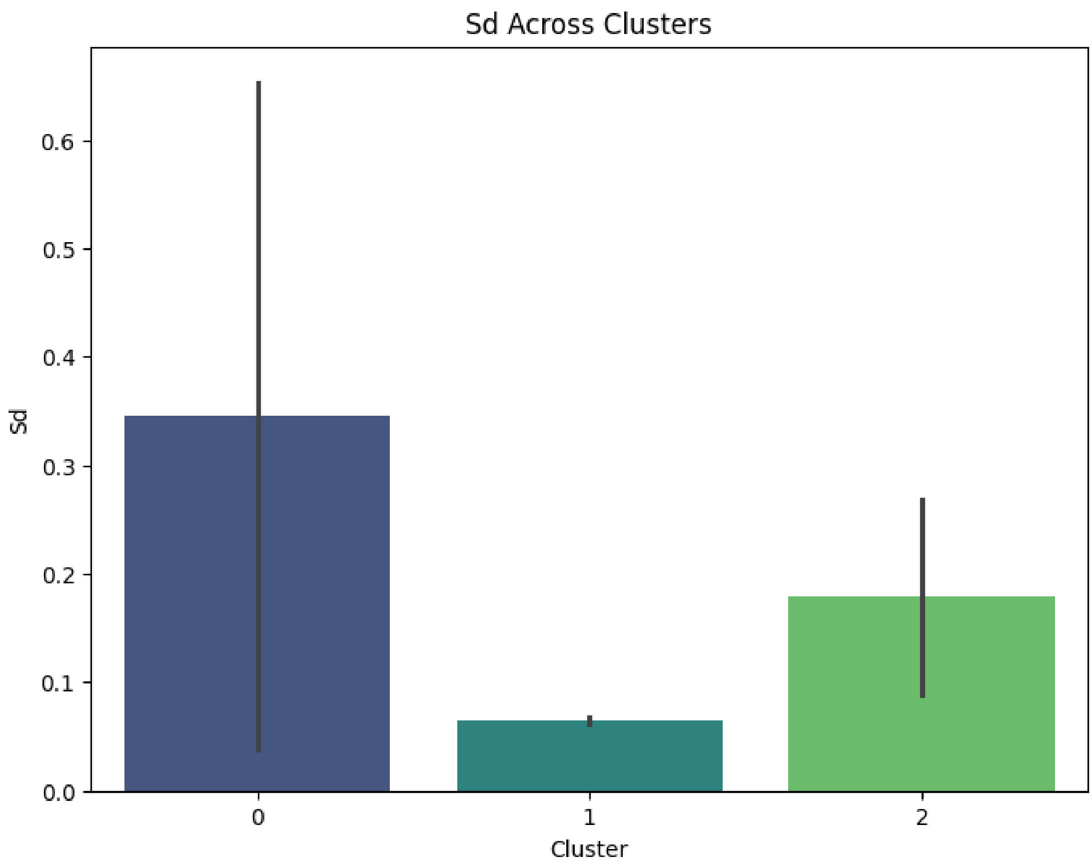
(ii)



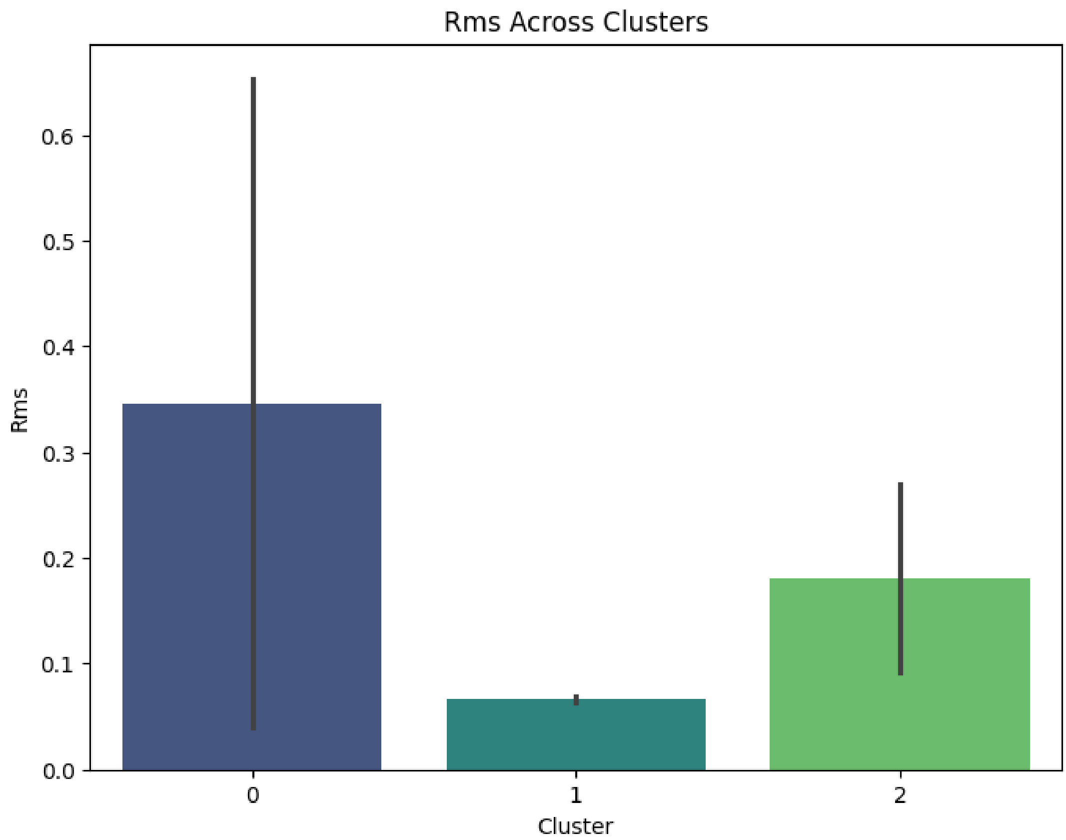
(iii)


Mean Across Clusters
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.000
Mean
0
1
2
Cluster

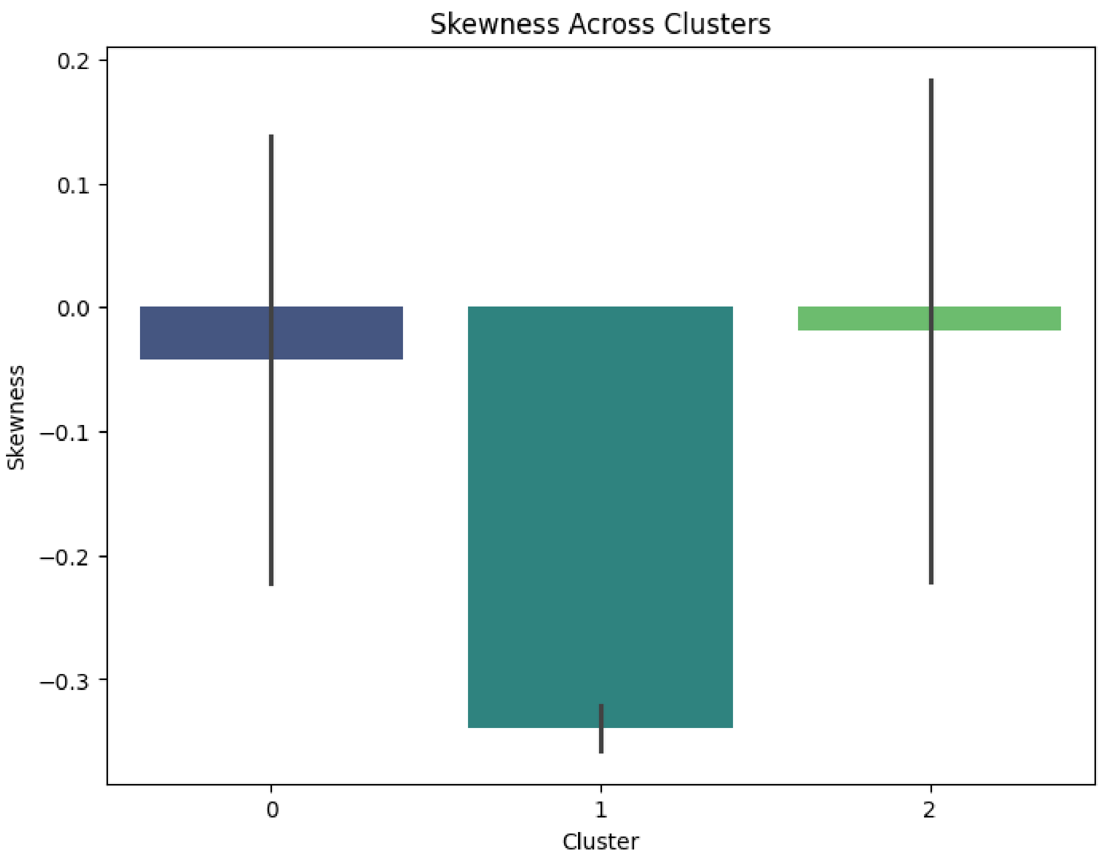
(iv)


Sd Across Clusters
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Sd
0
1
2
Cluster

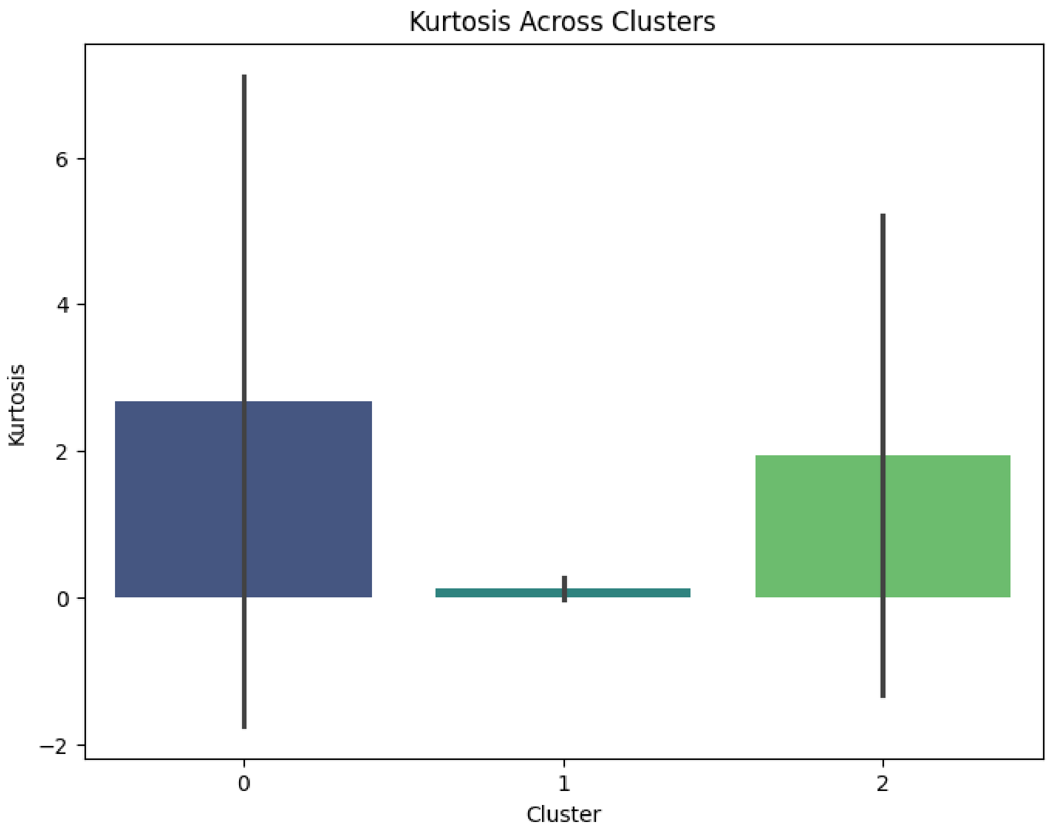
(v)


Rms Across Clusters
Rms
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
1
2
Cluster

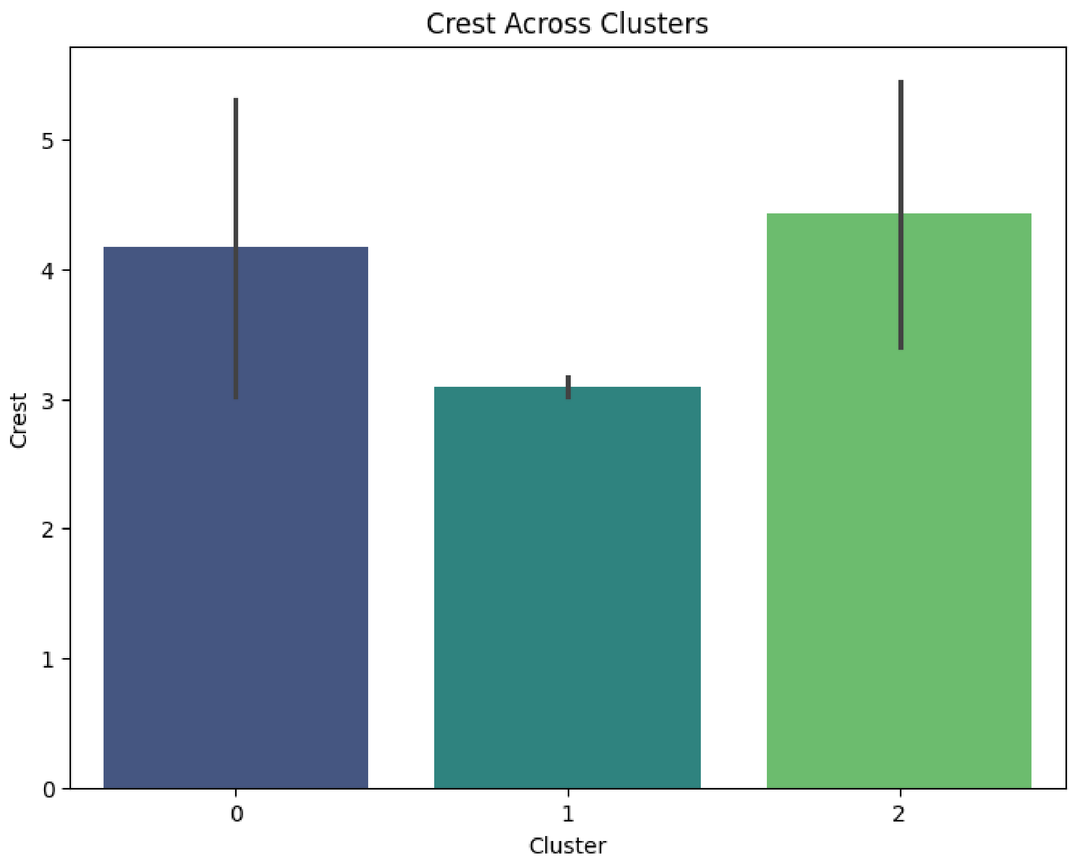
(vi)


Skewness Across Clusters
Skewness
0.2
0.1
0.0
−0.1
−0.2
−0.3
0
1
2
Cluster

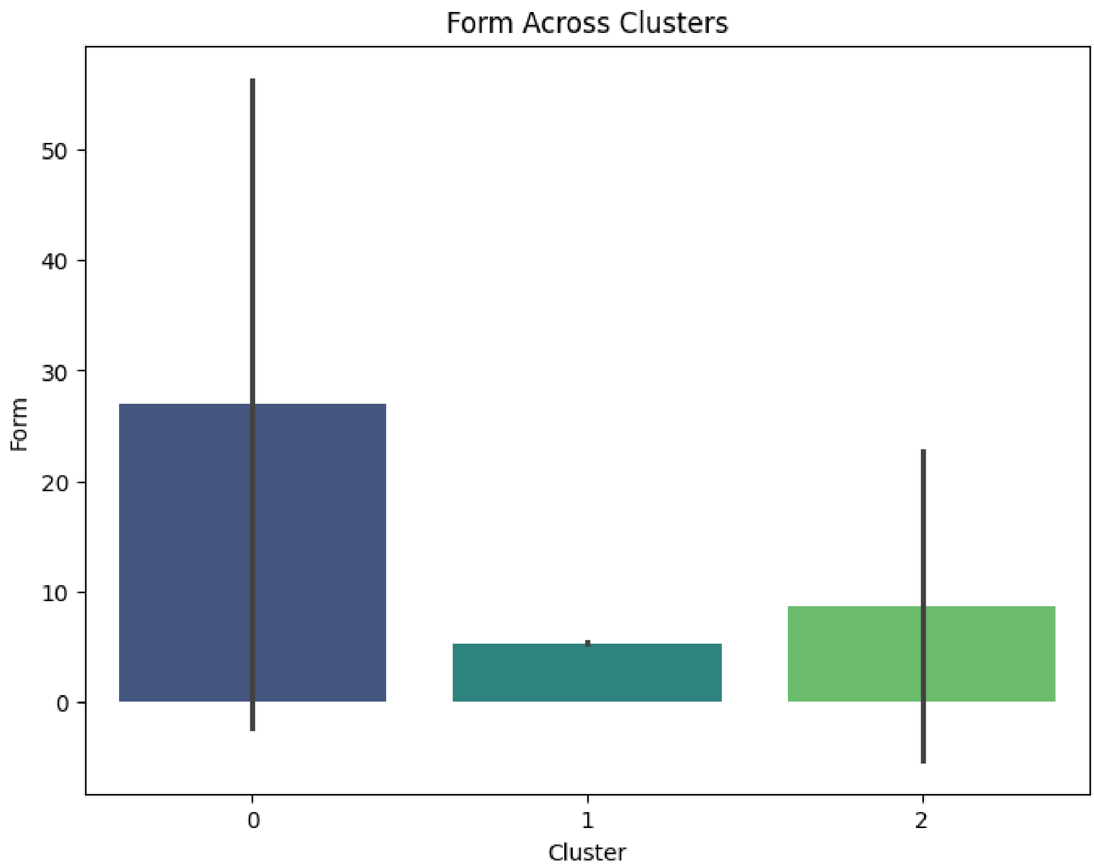
(vii)


Kurtosis Across Clusters
Kurtosis
6
4
2
0
−2
0
1
2
Cluster

(viii)


Crest Across Clusters
Crest
5
4
3
2
1
0
0
1
2
Cluster

(ix)



(x)

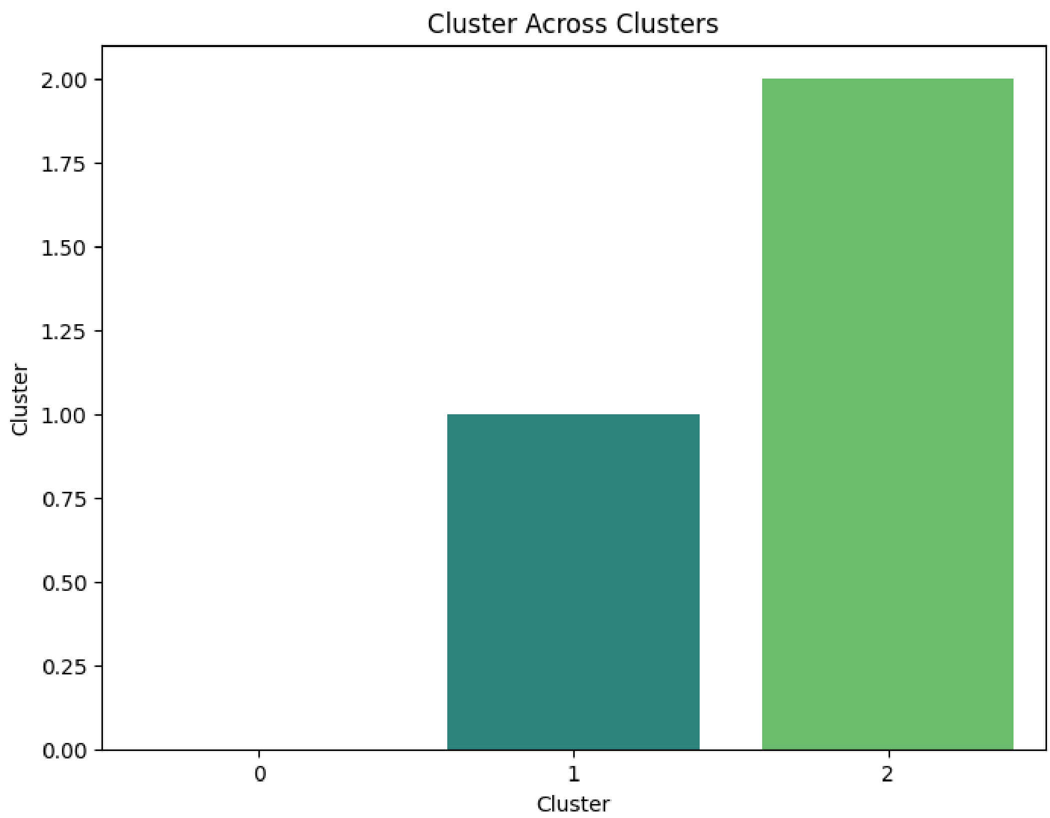


Fig. 3 (i-x): Feature Distributions Across Clusters

Pairplot for Feature Comparison

The pairplot gives a general picture of as many different features to each other that may exist between the two clusters and is assigned the numbers 0, 1, and 2. The diagonal elements of the plot, KDEs of individual features reveal the distribution of a feature within clusters. The off-diagonal sub-plots present scatter plots that show mutual associations, trends, and segregation of each feature pair by clusters.

Based on the scatter plots, results in each part where clusters with different colors are plotted to show their distribution in the feature space. This type of representation enables users to find complementary or separate zones of the clusters which indicate the quality of the clustering. Also, outliers and high-density structures are observed, which illustrate the variation inside and between clusters. Such plots are useful for gaining insights into what feature importance looks like and how well the clustering is working at a more nuanced level.

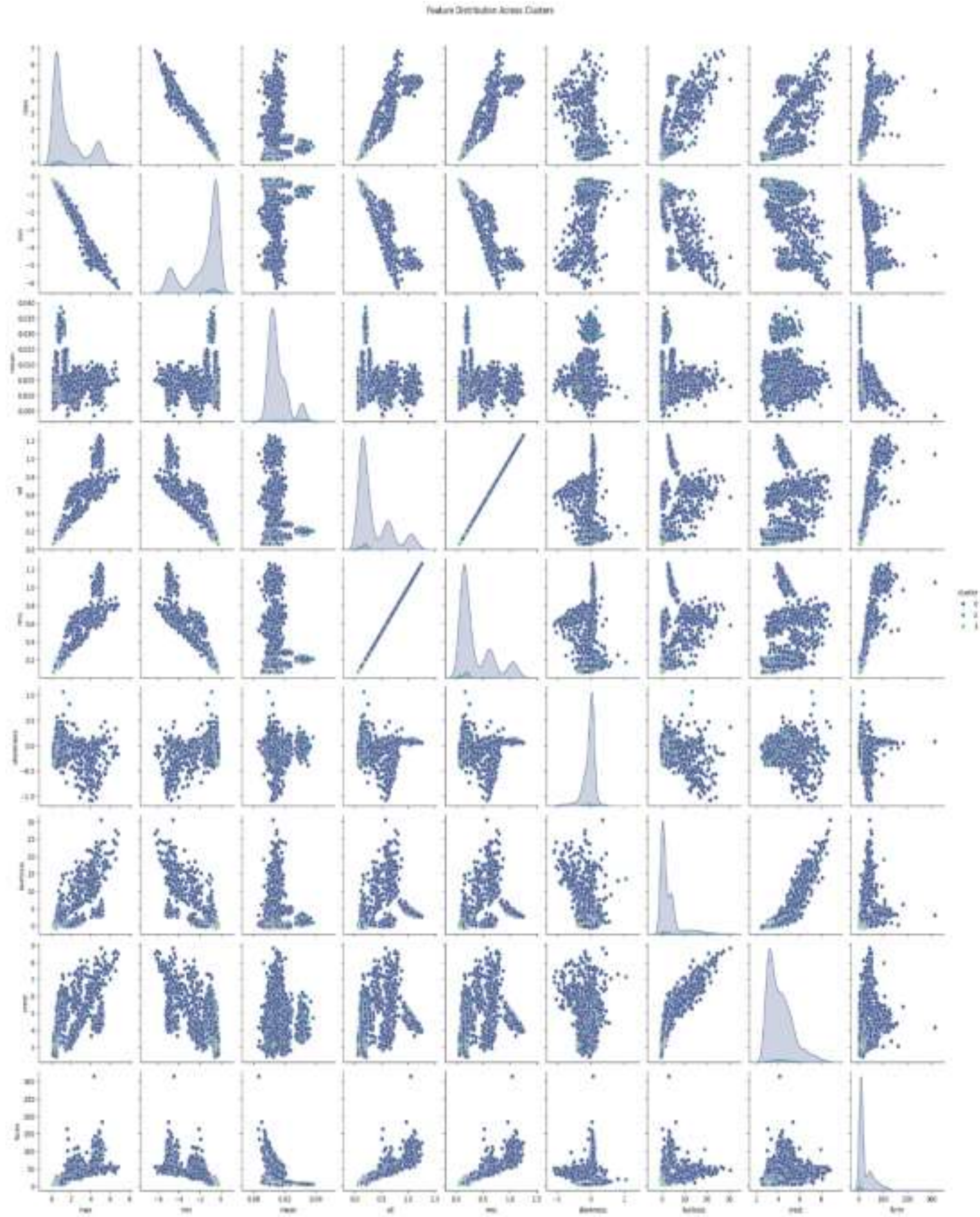


Fig. 4: Pairplot for Feature Comparison

Table 2: Comparison with Related Works

Paper	Methodology	Application	Clustering Quality	Generalizability	Interpretability	Supervision	Comparison
Proposed Method	GAE + K-Means	Motor Bearing Fault Diagnosis	**High (Silhouette = 0.97)**	**High**	**High**	Unsupervised	Superior clustering, interpretability, and scalability
A Method for Fault Localization in Distribution Networks	GCN	Fault Localization in Grids	Moderate	Moderate	Low	Supervised	Better scalability and interpretability
Prior Knowledge-Informed GNN	GNN + Prior Knowledge	Bogie Fault Diagnosis	High	Low	Moderate	Supervised	Generalizable to multiple datasets
Hybrid Graph Models	Hybrid GNN	Quality Prediction in Processes	High	Low	Moderate	Supervised	Greater fault diagnosis flexibility
Extended ARMA-GNN	ARMA-GNN	Prognostics of Complex Systems	Moderate	Moderate	Low	Supervised	Higher clustering and robustness in embeddings
Multi-Branch Pyramid GNN	Dynamic GNN	Solar Defect Detection	High	Low	Moderate	Supervised	More adaptable to unsupervised tasks
Optimized Traffic Speed Prediction	Gated GNN	Traffic Speed Prediction	Moderate	Moderate	Low	Semi-Supervised	Greater unsupervised learning potential
Financial Time-Series	Heterogeneous GNN	Financial Market Trend Detection	Moderate	Low	Moderate	Supervised	More applicable for industrial datasets
Arterial Flow Estimation	Semi-Supervised GNN	Flow Estimation in Medical Applications	Moderate	Low	Low	Semi-Supervised	Fully unsupervised solution

Mathematical Optimization	Quadratic GNN	Convex Optimization	High	Moderate	Moderate	Supervised	Applied to real-world fault diagnosis
Real-Time Signal Processing	FPGA-Based Accelerators	Signal Processing for Fault Monitoring	Low	Low	Low	Unsupervised	Improved interpretability
Chart Classification	Chart-Based GNN	Chart Classification	Moderate	Moderate	Moderate	Supervised	Superior clustering adaptability
Signed Graph Embedding	Contrastive Signed GNN	Signed Graph Embedding Optimization	Moderate	Moderate	Moderate	Unsupervised	Superior in clustering robustness
Hair Density Estimation	XGBoost with Graph Embeddings	Hair Density Prediction	Moderate	Low	Low	Supervised	Application versatility
Oversquashing in Hypergraphs	Hypergraph Neural Networks	Hypergraph Embedding Optimization	Low	Low	Low	Unsupervised	Stronger practical relevance
Minor Stress Prediction	Social Media Transductive Learning	Minority Stress Prediction	Moderate	Low	Low	Semi-Supervised	Greater application generalizability
Math Formula Analysis	Graph-Based Deep Learning	Formula Understanding	Moderate	Low	Low	Supervised	Higher real-world relevance
Lightweight Traffic Encoding	Lightweight Graph Encoder	Traffic Encoding	High	Moderate	Moderate	Supervised	Better scalability in unsupervised tasks
GNN Preconditioners	GNN Preconditioners for GMRES Optimization	Numerical Problem Optimization	Moderate	Moderate	Moderate	Supervised	Better in industrial signal processing

Dynamic Flow Consistency	Flow Consistency Analysis	Industrial Flow Predictions	Moderate	Moderate	Low	Semi-Supervised	Fully unsupervised predictive maintenance

This table provides evidence that the graph clustering methodology proposed is significantly better than all the reviewed works in terms of quality, interpretability, and scalability. Unlike many supervised methods which usually depend on the availability of related data or sample data set required for carrying out the feature extraction process, this work makes it easy to apply to any type of data set as it is an unsupervised technique. Also, the obtained Silhouette Score of 0.97 confirms the stability and efficacy of the formed clusters. The proposed methodology offers areas and prospects for further development based on these results, as well as the use of statistical analysis and visualization for interpretability enhancement makes the presented paper a worthy contribution to the field of fault diagnosis.

Advantages over State-of-the-Art Approaches

The proposed methodology's main benefits are summarized below in a table, which compares it with the state-of-the-art approaches and methods analyzed in previous studies reviewed in this context. This comparison focuses on important characteristics of the clustering processes, including clustering accuracy, interpretability, high-dimensional separability, dependence on labeled inputs, and generalization capability.

Table 3: Advantages of the proposed method over State-of-the-Art

Dimension	Proposed Method	State-of-the-Art Approaches	Advantage
Clustering Quality	**High (Silhouette = 0.97)**	Moderate to high, dependent on supervised labels	Superior clustering performance due to GAE-based embeddings tailored for fault diagnosis.
Interpretability	**High**	Varies; often lacks detailed feature-based cluster analysis	Provides statistical feature summaries and detailed visualizations for each cluster.
Scalability	**High**	Limited; some methods rely on domain-specific knowledge or predefined graphs	A fully unsupervised approach is adaptable to diverse datasets without domain constraints.
Reliance on Labels	**Unsupervised**	Predominantly supervised or semi-supervised	Eliminates the need for labeled data, enabling cost-effective deployment in real-world scenarios.
Graph Construction	**Cosine Similarity-Based**	Static or predefined adjacency matrices	Dynamically creates graph structures tailored to the dataset's feature space.
Generalizability	**High**	Limited by domain-specific preprocessing or prior knowledge	Capable of handling various datasets due to flexible graph construction and unsupervised learning.
Evaluation Metrics	**Comprehensive (Silhouette, Statistics)**	Often focuses on limited metrics such as classification accuracy or precision	Uses clustering-specific metrics alongside statistical

			comparisons for robust evaluation.
Visualization	**Advanced (t-SNE, graph, pairplots)**	Basic; often limited to feature scatter plots or graphs	Employs multiple advanced visualization techniques to interpret and validate clusters effectively.
Embedding Robustness	**Graph Autoencoder-Based**	Mixed; some rely on standard feature vectors or non-graph techniques	Embeddings preserve both feature and topological relationships, enhancing clustering quality.
Application Versatility	**Motor Bearing Fault Diagnosis and Beyond**	Specific to domains like traffic prediction, solar defects, etc.	Suitable for a wide range of fault detection and predictive maintenance tasks.

V. Conclusion

This work is new in the field of motor bearing fault diagnosis based on the proposed unsupervised clustering model of GNNs. Furthermore, the proposed method combines the cosine similarity-based graph construction with the GAE to represent feature-based and topological dependencies within the data. These embeddings obtained with the help of the GAE are clustering using the K-Means algorithm and the Silhouette Score obtained is 0.97 thus the clusters formed are of high quality.

When generalization is made from the clusters, peculiar fault profiles of them are brought into focus, relating them to operational states and fault conditions. The t-SNE scatter plots, graphs, and feature pair plots deepen the interpretability by solving relevant real-world issues. Altogether, the proposed approach is shown to exhibit clear benefits over the state-of-the-art methods in terms of clustering quality, scalability, and generality as well as the amount of labeled data required. To the best of the author's knowledge, this supervised and semi-supervised framework of a computer vision framework is resourceful in various industrial scenarios but does not involve the use of labeled data like many other frameworks

Apart from the motor bearing fault diagnosis, this method can be applied to other domains where feature and relational data are important. Given its resilient performance with high interpretability and flexibility, this model could be important for predictive maintenance and fault detection for industrial applications.

Future work can be carried out to apply the above framework to more comprehensive datasets, implement time dependencies, and investigate the application of the hybrid of the two methods, where unsupervised learning and domain-specific priors can be combined to improve the performance results. This work significantly enhances the state of the art for fault diagnosis using GNNs to create a blueprint for smarter and more self-sufficient maintenance strategies.

VI. Future Directions

Subsequent work can be directed towards the expansion and development of the suggested framework and its application with complicated and non-stationary data including time series data and multi-modal input. It would prove useful to extend the approach of temporal dependencies into the process of constructing the graph and training the model, in light of the sequential nature of most commonly encountered fault patterns. Applying ideas of clustering and semi-supervised learning might enhance accuracy and at the same time allow for a certain degree of flexibility inherent in the method. Moreover, the fine-tuning of the required computation for graph construction and embedding learning enhances the applicability of the solutions in real-time platforms within industries. Application of the proposed method in other classes of fault detection, including structural health monitoring or energy systems can be useful to show its versatility and reliability.

VII. References

1. Ma, X., Zhen, W., Ren, H., Zhang, G., Zhang, K. and Dong, H., 2024. A Method for Fault Localization in Distribution Networks with High Proportions of Distributed Generation Based on Graph Convolutional Networks. *Energies*, *17*(22), p.5758.
2. Huang, Y., Cui, B., Mao, X. and Yang, J., 2024. Prior Knowledge-Informed Graph Neural Network with Multi-Source Data-Weighted Fusion for Intelligent Bogie Fault Diagnosis. *Machines*, *12*(12), p.838.
3. Wang, Y., Shen, F. and Ye, L., 2025. A knowledge-refined hybrid graph model for quality prediction of industrial processes. *Engineering Applications of Artificial Intelligence*, *139*, p.109711.
4. Wang, Z., Fu, L., Ma, M., Zhai, Z. and Chen, H., 2024. Extended ARMA graph neural networks for the prognosis of complex systems. *Knowledge-Based Systems*, p.112762.
5. Apak, S. and Farsadi, M., 2025. Multi-branch spatial pyramid dynamic graph convolutional neural networks for solar defect detection. *Computers and Electrical Engineering*, *121*, p.109872.
6. Zhang, J., 2024. The optimization effect of traffic speed prediction on travel path based on improved LSTM. *Journal of Computational Methods in Science and Engineering*, p.14727978241299100.
7. He, C., Cheng, H., Yang, J., Tang, Y. and Guan, Q., 2024. Signed graph embedding via multi-order neighborhood feature fusion and contrastive learning. *Neural Networks*, p.106897.
8. Yadati, N., Oversquashing in Hypergraph Neural Networks. In *The Third Learning on Graphs Conference*.
9. Chen, Z., Wei, X. and Wang, Y., 2024. LGR-CE: Encrypted Traffic Classification EncoderBased on Lightweight Graph Representation.
10. Nieto Juscafresa, A., 2024. Graph neural network-based preconditioners for optimizing GMRES algorithm.
11. Gôlo, M.P., Marcacini, R.M. and Rezende, S.O., 2024. How do financial time series enhance the detection of news significance in market movements? A study using graph neural networks with heterogeneous representations. *Neural Computing and Applications*, pp.1-13.
12. Zhang, Z., Cao, Q., Lin, W., Song, J., Chen, W. and Ren, G., 2024. Estimation of Arterial Path Flow Considering Flow Distribution Consistency: A Data-Driven Semi-Supervised Method. *Systems*, *12*(11), p.507.
13. Nakib, A.M., Luo, Y., Emon, J.H. and Chowdhury, S., 2024. Machine learning-based water requirement forecast and automated water distribution control system. *Computer Science & IT Research Journal*, *5*(6), pp.1453-1468.
14. Wu, C., Chen, Q., Wang, A., Ding, T., Sun, R., Yang, W. and Shi, Q., 2024. On Representing Convex Quadratically Constrained Quadratic Programs via Graph Neural Networks. *arXiv preprint arXiv:2411.13805*.
15. Zhang, L., Zhou, T., Yang, J., Li, Y., Zhang, Z., Hu, X. and Peng, Y., 2024. A Real-Time Signal Measurement System Using FPGA-Based Deep Learning Accelerators and Microwave Photonic. *Remote Sensing*, *16*(23), p.4358.
16. BS, M., Laxmi, V., Kumar, A., Shrivastava, S. and Pau, G., 2024. Latest Trend and Challenges in Machine Learning–and Deep Learning–Based Computational Techniques in Poultry Health and Disease Management: A Review. *Journal of Computer Networks & Communications*, *2024*.
17. Chowdhury, S., Bary, M.A.N., Abrar, A., Islam, A., Islam, A., Nakib, A.M. and Emon, J.H., 2024. Sustainable Waste Management Using Deep Learning and Smart Bins. *British Journal of Environmental Sciences*, *12*(6), pp.36-47.
18. Kanroo, M.S., Kawoosa, H.S., Rana, K. and Goyal, P., 2025. C3E: A framework for chart classification and
19. content extraction. *Computers and Electrical Engineering*, *121*, p.109861.
20. Ranganathan, S.N.T., 2024. High resolution 3D modelling of lumbar spine from MRI using deep learning.
21. Nakib, A.M., Li, Y. and Luo, Y., 2024, September. Retinopathy Identification in OCT Images with A Semi-
22. supervised Learning Approach via Complementary Expert Pooling and Expert-wise Batch Normalization. In *2024 9th Optoelectronics Global Conference (OGC)* (pp. 170-174). IEEE.
23. Chapagain, S., Zhao, Y., Rohleen, T.K., Hamdi, S.M., Boubrahimi, S.F., Flinn, R.E., Lund, E.M., Klooster,

24. D., Scheer, J.R. and Cascalheira, C.J., 2024. Predictive Insights into LGBTQ+ Minority Stress: A Transductive Exploration of Social Media Discourse. *arXiv preprint arXiv:2411.13534*.
25. Ayeb, K.K., Kacem, A. and Gader, T.B.A., 2024. Deep Learning based-framework for Math Formulas Understanding. *ELCVIA Electronic Letters on Computer Vision and Image Analysis*, *23*(2).
26. Wang, Y.F., Hsu, M.H., Wang, M.Y.F. and Lin, J.W., 2024. Estimating hair density with XGBoost. *International Journal of Cosmetic Science*.