

# A dynamic graph structure optimization diagnosis

Zhiyuan Hu<sup>1</sup>, Yangde Lin<sup>1</sup>, Jianrong Li<sup>1</sup>(✉) and Juan Lyu<sup>1</sup>

<sup>1</sup> College of Artificial Intelligence, Tianjin University of Science and Technology Tianjin,  
China  
lisa\_ljr@tust.edu.cn

**Abstract.** In the field of industrial equipment management and academic network analysis, early fault diagnosis and node classification tasks are of great significance for ensuring the stable operation of equipment and promoting knowledge discovery. Existing methods face many challenges in dealing with large-scale and unbalanced data sets, especially in bearing fault diagnosis and scientific literature classification. In response to these challenges, this paper proposes a Dynamic Graph-Structured Optimization Diagnosis model based on graph neural network. The innovation of the model primarily encompasses two aspects. Firstly, concerning the dataset, the k-nearest neighbor algorithm is utilized to fuse the health status of bearings with vibration signal data. This integration facilitates the construction of a graph structure that accurately captures the complex relationship between different bearing states. At the same time, an optimization strategy combining Focal Loss and graph Deep Open Classification method is used to further improve the applicability and accuracy in different fields on the basis of enhancing the performance of the model in dealing with unbalanced data. During the experiment, the DG-SOD model showed excellent performance in the above tasks. The accuracy of bearing fault diagnosis increased to 65%, the accuracy of Core node classification increased from 76 % to 86.65 %, and the classification accuracy of CiteSeer increased from 70 % to 76.05 %. The above data show that the DG-SOD model has obvious advantages in dealing with data imbalance problems in industrial equipment detection and scientific literature classification and improving the accuracy of minority class recognition. It provides new ideas and frameworks for future industrial equipment management and academic network analysis.

**Keywords:** bearing fault diagnosis, graph neural network, k nearest neighbor algorithm, Focal Loss.

## 1 Introduction

With the continuous improvement of industrial automation and intelligence, the efficiency and reliability of equipment management have become particularly important. As a key component widely used in various mechanical equipment, the health status of bearings is directly related to the stable operation and production safety of equipment. Therefore, the development of effective bearing fault early diagnosis methods is of great significance for preventing equipment failure and reducing maintenance costs. Traditional fault diagnosis methods, such as handheld device detection and temperature

measurement, have been applied in practice, but these methods rely on a high degree of manual experience, lack of real-time and accuracy, and are difficult to meet the needs of modern industry for high efficiency and high reliability.

At the same time, in the academic field, how to effectively classify and analyze a large number of scientific literatures to promote the discovery and dissemination of knowledge is also a long-standing challenge. Especially in today's rapid development of information technology, the number of scientific publications is growing exponentially, which brings great pressure to the management and analysis of literature.

In response to the above challenges, graph neural network (GNN) [1] has long been widely concerned as an effective data analysis tool. GNN can capture complex relationships and structural information in data, and shows superior performance on multiple tasks, including node classification, link prediction and graph classification. However, current challenges faced by GNN in large-scale [2] and imbalanced datasets, such as bearing fault diagnosis and scientific literature classification, include the ability to handle dynamic data changes, address class imbalance issues, and effectively identify new categories.

To solve these problems, this study proposes a novel method, dynamic graph structure optimization diagnosis (DG-SOD) model. The model innovatively uses the k-nearest neighbor algorithm [3] to construct a graph structure based on bearing health status and vibration signals, and introduces an improved Focal Loss function [4] and an optimization strategy based on the graph Deep Open Classification (gDOC) [5] method. The combination of these technologies not only optimizes the model's ability to deal with data imbalance, but also improves the applicability and accuracy of the model in many fields such as bearing fault diagnosis and scientific literature classification.

The paper presents two significant contributions. Firstly, a novel DG-SOD model is introduced, integrating the k-nearest neighbor algorithm to construct a graph structure. This integration effectively captures intricate relationships within bearing states and vibration signal data, enhancing the model's adaptability to dynamic data changes and extending its utility and accuracy across diverse domains. Secondly, an optimization strategy is proposed that combines Focal Loss with the graph Deep Open Classification (gDOC) method, notably boosting the efficacy of the model in handling imbalanced datasets and identifying new categories.

The subsequent sections of this paper are organized as follows. In Section 2, related works in industrial equipment management and academic network analysis are reviewed, with a focus on challenges in early fault diagnosis and node classification. Section 3 introduces the DG-SOD model, elucidating its components such as the graph structure constructed via the k-nearest neighbor algorithm, and the integration of Focal Loss and gDOC methods. Section 4 delineates the experimental setup, detailing dataset specifics, preprocessing procedures, parameter configurations, and baseline models used for comparison. Experimental findings are presented in Section 5, where the performance of the DG-SOD model is analyzed across various datasets, accompanied by discussions on potential enhancements. Finally, Section 6 explores the implications of the study's outcomes, potential applications of the DG-SOD model, and avenues for

future research. This section culminates by emphasizing the model's significance in advancing both industrial equipment management and academic network analysis.

## 2 Related works

In the context of industrial equipment management and academic network analysis, early fault diagnosis and node classification tasks are two important research directions. Bearing fault diagnosis is the key to ensure the stable operation of mechanical equipment. Traditional fault diagnosis methods rely on experienced technicians to judge by listening and feeling. These methods are not only time-consuming and labor-intensive, but also difficult to ensure accuracy and real-time performance. In recent years, with the combination of vibration analysis technology, big data, artificial intelligence and other technologies, by analyzing the vibration signal of the bearing, the mechanism + data-driven method can effectively improve the accuracy and real-time performance of fault diagnosis, and the research of bearing fault diagnosis has made remarkable progress [6].

Academic network analysis mainly involves tasks such as classification and clustering of scientific literature [7], which aims to promote the discovery and dissemination of knowledge. However, traditional literature management and analysis methods are difficult to cope with the exponential growth of literature under the background of rapid development of information technology. As an effective data analysis tool, graph neural network (GNN) can capture complex relationships and structural information in data, and shows superior performance in many tasks such as node classification, link prediction and graph classification. However, when GNN is directly applied to bearing fault diagnosis [8] and scientific literature classification, it still faces the challenge of low accuracy when dealing with large-scale and unbalanced data sets.

Over time, the structure and content of graph data may change, such as the addition of new nodes, the removal of old nodes, or changes in the relationship between nodes. These dynamic changes bring new challenges to graph data analysis. Lifelong learning aims to enable the model to accumulate knowledge in the process of continuous learning, reduce the forgetting of old tasks, and improve the adaptability to new tasks. In the existing research, although some solutions have been proposed, such as incremental learning strategy [9] and experience playback method [10], there are still some limitations in dealing with unbalanced class problems and new class recognition.

Although some progress has been made in the field of bearing fault diagnosis and academic network analysis, there are still challenges in dealing with large-scale, unbalanced data sets and dynamic changes of adaptation graph data. The dynamic graph structure optimization diagnosis (DG-SOD) model proposed in this paper effectively responds to these challenges by innovatively combining k-nearest neighbor algorithm, improved Focal Loss function and optimization strategy based on gDOC method, and provides new ideas for future research and practice.

### 3 Proposed method

In this study, a dynamic graph structure optimization diagnosis (DG-SOD) model is introduced, along with a lifelong learning method, to address challenges in bearing fault diagnosis and academic network analysis, particularly focusing on efficient solutions to data dynamic changes and class imbalance problems. Figure 1 is an example of a lifelong graph learning problem for imbalanced classes. The key components and technological innovations of the model are introduced in detail below.

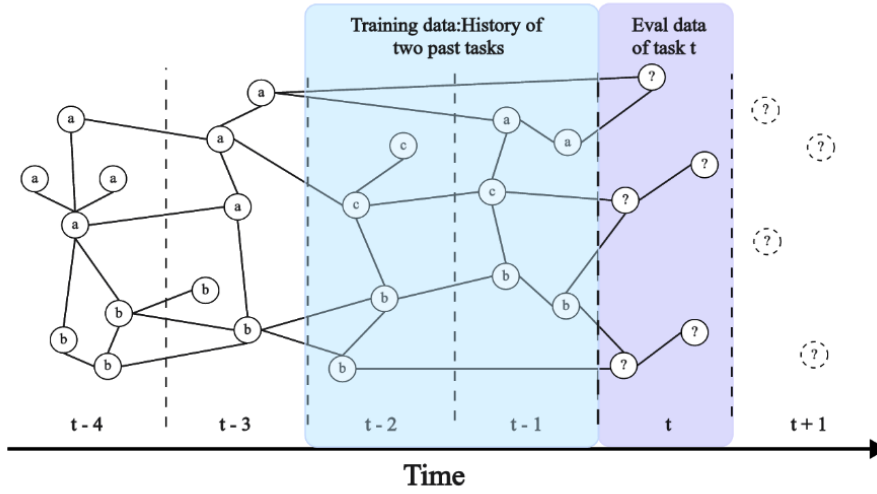


Fig. 1. Lifelong Graph Learning: Class Imbalance and New Class Detection

The problem of lifelong graph learning with class imbalance and new classes is illustrated. At time  $t$ , the learner has to classify new vertices of task  $T_t$  (purple). Any task might come with previously unseen classes. For example, the class “c” emerged only at task  $t-2$  and was subsequently added to the class set. The learner may utilize internal and external knowledge from previous tasks to adapt to the current task. Upon evaluating task  $T_t$ , the learning process proceeds to task  $T_{t+1}$ .

#### 3.1 Overview of the model

The dynamic graph structure optimization diagnosis (DG-SOD) model combines the deep learning ability of graph neural network (GNN) and advanced optimization techniques for dealing with data imbalance problems. The DG-SOD model is especially suitable for dynamically changing data sets, such as time series data or continuously updated network data. By integrating new information in real time and using KNN algorithm to dynamically update the graph structure [11], the model can adapt to changes

in data and ensure that the classification accuracy is not affected by time changes. The combination of Focal Loss function and gDOC method further enhances the model's ability to deal with data imbalance problems, making it perform well in new category detection.

In the face of emerging research topics or unknown fault types, the DG-SOD model can quickly adapt to and accurately identify new categories through its optimized learning strategies. This ability is of great significance for early fault diagnosis and real-time tracking of academic trends.

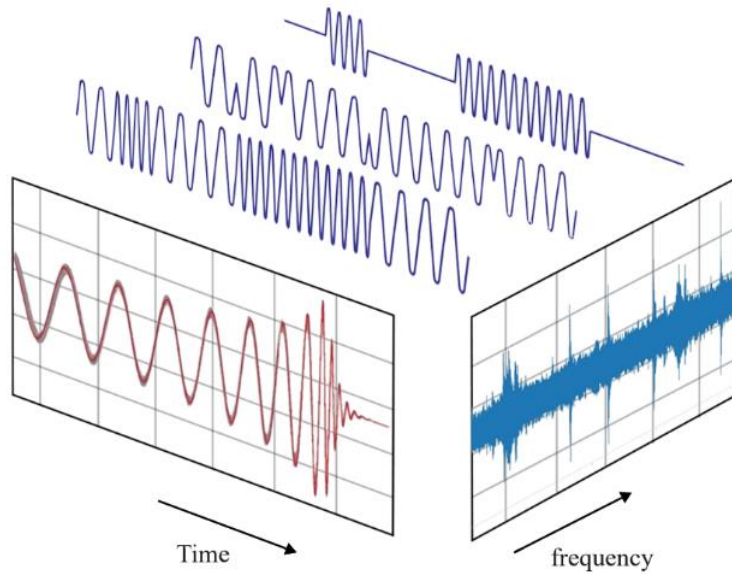
### 3.2 Graph structure construction

The bearing fault data set is specially designed for bearing health monitoring and fault diagnosis [12], which contains a variety of bearing vibration signal samples under different operating conditions. These samples cover conditions from normal operation to a variety of fault states, such as inner ring, outer ring and rolling element faults. Each sample in the data set is a vibration signal collected by a high-speed sensor under actual working conditions. These signals are preprocessed and used to train and test the fault diagnosis model. The data set usually contains thousands of samples. Each sample consists of a continuous vibration signal and a corresponding state. The sampling rate and signal length are determined according to the specific experimental design. This data set is of great significance for improving the reliability of mechanical equipment and implementing preventive maintenance strategies.

The Cora dataset is an open graph-structured dataset, which is widely used in the research of graph neural networks and machine learning. It consists of about 2,700 academic papers and 5429 edges composed of citations between them, each of which is marked with one or more category labels. The node represents the paper, and the edge represents the citation relationship between the papers, forming a complex academic citation network. The main purpose of the Cora dataset is to use it for graph-structured node classification tasks, especially in the fields of recommendation systems and natural language processing. Each paper in the data set has a feature vector. These feature vectors are usually extracted from the abstract of the paper by the TF-IDF [13] method, which provides the content information of the nodes for the graph neural network.

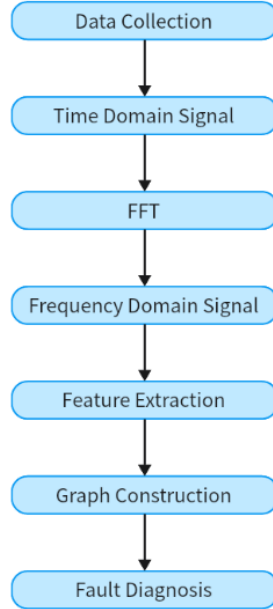
The CiteSeer dataset is an academic network dataset similar to Cora, but it covers a wider range of research areas and more academic papers. The dataset contains about 3,300 papers and more than 43,000 citations, and each paper is also assigned a category. The node represents the paper, and the edge represents the citation relationship between the papers, which constitutes a huge academic citation graph. The CiteSeer dataset is widely used in graph-structured node classification tasks, especially in the fields of academic network analysis and information retrieval. The feature vectors of the papers in the dataset usually include metadata such as the title, keywords and abstracts of the papers, which are helpful for the graph neural network model to better understand and classify academic literature.

For the bearing fault dataset, it comprises vibration signals collected from industrial bearings, which directly influence the stability and safety of mechanical equipment. To accurately extract the intricate features embedded in the signal, the fast Fourier transform (FFT) [14] is initially employed to convert the time-domain signal into a frequency-domain signal, as depicted in Figure 2. This procedure unveils the spectral characteristics of the signal, enabling the identification of specific frequency components associated with bearing faults.



**Fig. 2.** Transform the time domain signal into a spectrum signal

Then, based on the frequency domain signal, we use the k-nearest neighbor algorithm (KNN) to construct a complex graph structure. In Fig.3, the nodes represent a single vibration signal sample, and the edges between the nodes are established according to the similarity of frequency domain features. Through this method, the different operating states of the bearing are effectively mapped to the graph structure, which provides rich structured information for subsequent fault diagnosis.



**Fig.3.** Flowchart of Bearing Fault Diagnosis Process

### 3.3 Applications of GNN combined with lifelong learning framework

The GNN architecture suitable for processing graph data is selected and adjusted, allowing the model to capture the structural features between nodes by aggregating neighbor node information. The model uses advanced graph neural network techniques to adapt to complex data patterns including dynamic graph structures [15].

The lifelong learning framework of the DG-SOD model adopts an iterative up-date strategy [16], so that the model retains the learned knowledge while learning new data, effectively avoids catastrophic forgetting, and ensures the stability and reliability of the model in the continuous learning process.

### 3.4 Loss function and optimization strategy

To address the common classification challenges encountered in highly imbalanced datasets, an innovative approach is adopted, which combines the Focal Loss function with the gDOC strategy. The Focal Loss function has been proved to be particularly effective in dealing with imbalanced data sets because it can increase the attention of the model to hard-to-classify samples [17]. By adjusting the loss function, it reduces the weight of easy-to-classify samples, so that the model pays more attention to those minority samples that are difficult to identify in the training process. The Focal loss form is as follows.

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log p_t \quad (1)$$

Among them,  $p_t$  is the probability that the model is predicted to be a positive category, and  $\alpha_t$  is an adjustable parameter used to adjust the importance weight of positive and negative samples.  $\gamma$  is a factor that controls the adjustment of difficult and easy samples. When  $\gamma > 0$ , it will reduce the loss of easy samples, thus focusing more attention on difficult samples.

The gDOC strategy aims to provide a flexible category detection mechanism that can not only deal with the problem of unbalanced category distribution, but also identify and adapt to emerging research topics or unknown fault types in a timely manner.

In the actual implementation, the combination of the improved Focal Loss function and the gDOC method now adjusts the loss calculation. For each sample, its difficulty and category attribution are initially evaluated through the gDOC method, followed by the adjustment of focus parameters in the Focal Loss based on this evaluation. This dynamic adjustment mechanism ensures that the model can adaptively adjust the attention to different samples during the training process, especially for those samples with a small number of categories or emerging categories.

## 4 Experiment

### 4.1 Baseline model

To comprehensively evaluate the performance of the DG-SOD model, three graph neural networks were selected as baseline models for comparison. These include the Graph Convolutional Network (GCN), which captures local structural information of nodes through a neighbor aggregation mechanism [18], the Graph Attention Network (GAT), which introduces an attention mechanism allowing different weights to be assigned to different neighbors [19], and the Simple Graph Convolution (SGC), a simplified version of a graph convolution network that reduces computational complexity and model parameters [20].

### 4.2 Data preprocessing

From the vibration signal data of the bearing, a feature matrix is constructed [21]. Utilizing the k-nearest neighbor algorithm, similar neighbors are identified based on sample features, subsequently forming a complex graph structure. This process not only captures the relationship between samples, but also prepares the data base for the input of the deep learning model.

For Cora and CiteSeer datasets, the two datasets themselves provide rich graph structure information, where nodes represent academic papers and edges represent citation relationships between papers. In order to enhance the model's understanding of node content, we use the TF-IDF method to encode the text features of each node [22]. This



step effectively captures the importance of words in representing document topics by calculating the product of word frequency and inverse document frequency [23].

Furthermore, employing TF-IDF encoding [24], a distinctive feature vector is generated for each paper. This approach preserves the abundant information contained within the paper content while simultaneously reducing the complexity of the model for processing high-dimensional text data. This enables the graph neural network to learn and predict the classification of papers more effectively, and significantly improves the accuracy of node classification tasks.

$$TF(t, d) = \frac{f_{t,d}}{\sum_{k=1}^N f_{k,d}} \quad (2)$$

Among them,  $f_{t,d}$  is the number of times the word  $t$  appears in document  $d$ , and  $N$  is the total number of words in document  $D$ .

$$IDF(t, D) = \log \frac{N}{n_t} \quad (3)$$

Among them,  $N$  is the total number of documents in the corpus, and  $n_t$  is the number of documents containing the word  $t$ . In order to prevent the logarithm from being zero (that is, the words do not appear in the corpus), 1 can be added to the denominator.

$$TFIDF(t, d, D) = TF(t, d) * IDF(t, D) \quad (4)$$

Among them,  $TFIDF(t, d, D)$  is the TF-IDF value of word  $t$  in document  $d$ , and  $D$  is all documents in the corpus. The larger the value of TF-IDF, the more important the word is in the document.

### 4.3 Experimental parameter settings

In the experiments on all datasets, the following parameters were uniformly set. The dropout rate was fixed at 0.5 to prevent model overfitting. The length of historical information was set to 3, indicating that the model considered data from the first three time points. Two initial learning rates of 0.01 and 0.005 were tested during training. Additionally, the number of training epochs ranged from 200 to 500, adjusted based on the characteristics of different datasets.

### 4.4 Experimental results

We evaluated the performance of the dynamic graph structure optimization diagnosis (DG-SOD) model on three different datasets: bearing fault dataset, CiteSeer dataset, and Cora dataset. The following is the specific performance of each data set :

On the bearing fault data set, the accuracy of the DG-SOD model is 65 %. This shows that the model has certain recognition ability in complex fault diagnosis tasks [25], although there may still be room for improvement compared with traditional classification models.

The accuracy of the DG-SOD model on the CiteSeer dataset reached 76.05 %. This result shows the effectiveness of the model in the node classification task of the aca-

demographic literature citation network, especially when dealing with large-scale and unbalanced data sets, the model can effectively identify and classify different literature nodes.

On the Cora dataset, the accuracy of the DG-SOD model is 86.65 %, which shows the powerful performance of the model in the field of academic network analysis. The results demonstrate the excellent ability of the DG-SOD model in capturing graph structure features [26] and dealing with data imbalance problems, as well as its advantages in new category detection.

**Table 1.** Model performance index table under Cora data set

dataset	drop-out	history	initial_lr	annual_epochs	annual_lr	epoch	f1_macro	accuracy
Cora	0.5	3	0.01	250	0.01	250	0.8305	0.83759
Cora	0.5	3	0.01	250	0.01	250	0.8540	0.8665
Cora	0.5	3	0.005	200	0.005	200	0.8382	0.8422
Cora	0.5	3	0.005	200	0.005	200	0.8498	0.8633
Cora	0.5	3	0.01	200	0.01	200	0.8269	0.8329
Cora	0.5	3	0.01	200	0.01	200	0.8467	0.8623
Cora	0.5	3	0.01	250	0.01	250	0.8305	0.8376
Cora	0.5	3	0.01	250	0.01	250	0.8540	0.8665
Cora	0.5	3	0.005	250	0.005	250	0.8408	0.8469
Cora	0.5	3	0.005	250	0.005	250	0.8530	0.8665
Cora	0.5	3	0.01	250	0.01	250	0.8305	0.83759
Cora	0.5	3	0.01	250	0.01	250	0.8540	0.8665
Cora	0.5	3	0.005	300	0.005	300	0.8296	0.8364
Cora	0.5	3	0.005	300	0.005	300	0.8488	0.8633
Cora	0.5	3	0.01	250	0.01	250	0.8305	0.8376
Cora	0.5	3	0.01	250	0.01	250	0.8540	0.8665
Cora	0.5	3	0.01	250	0.01	250	0.8305	0.8376
Cora	0.5	3	0.01	250	0.01	250	<b>0.8540</b>	<b>0.8665</b>

**Table 2.** Model performance index table under CiteSeer data set

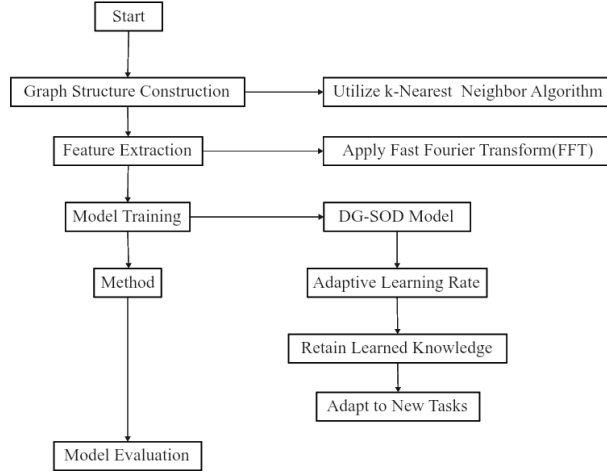
dataset	drop-out	history	initial_lr	annual_epochs	annual_lr	epoch	f1_macro	accuracy
CiteSeer	0.5	3	0.005	200	0.005	200	0.6704	0.7106
CiteSeer	0.5	3	0.005	200	0.005	200	0.7241	0.7579
CiteSeer	0.5	3	0.01	500	0.01	500	0.7198	0.7509
CiteSeer	0.5	3	0.01	500	0.01	500	0.6698	0.7069
CiteSeer	0.5	3	0.01	500	0.01	500	0.7198	0.7509

CiteSeer	0.5	3	0.005	1000	0.005	1000	0.6604	0.6976
CiteSeer	0.5	3	0.005	1000	0.005	1000	0.7265	0.7588
CiteSeer	0.5	3	0.01	250	0.01	250	0.6589	0.6994
CiteSeer	0.5	3	0.01	250	0.01	250	0.7270	0.7605
CiteSeer	0.5	3	0.02	250	0.02	250	0.6660	0.7059
CiteSeer	0.5	3	0.02	250	0.02	250	0.7223	0.7526
CiteSeer	0.5	3	0.01	250	0.01	250	0.6589	0.6994
CiteSeer	0.5	3	0.01	250	0.01	250	0.7270	0.7605
CiteSeer	0.5	3	0.01	220	0.01	220	0.6690	0.7069
CiteSeer	0.5	3	0.01	220	0.01	220	0.7177	0.7509
CiteSeer	0.5	3	0.01	240	0.01	240	0.6564	0.6948
CiteSeer	0.5	3	0.01	240	0.01	240	0.7103	0.7447
CiteSeer	0.5	3	0.01	250	0.01	250	0.6589	0.6994
CiteSeer	0.5	3	0.01	250	0.01	250	0.7270	0.7605
CiteSeer	0.5	3	0.01	250	0.01	250	0.6589	0.6994
CiteSeer	0.5	3	0.01	250	0.01	250	<b>0.7270</b>	<b>0.7605</b>

The parameters listed in the table include the dataset name, which represents the dataset used for the experiments. Dropout is a regularization technique utilized to prevent overfitting of neural networks. The term "History" refers to the length of historical information considered or the number of historical states used during model training. Initial\_lr denotes the initial learning rate, which determines the magnitude of weight updates at the start of model training. "Annual\_epochs" indicates the number of training cycles per year, while "Annual\_lr" represents the annual learning rate adjustment. In this study, the annual\_lr remains consistent with the initial\_lr, resulting in a constant learning rate throughout training. "Epoch" refers to the number of training iterations within a cycle. F1\_macro represents the macro-average of F1 scores, while "Accuracy" denotes the accuracy achieved by the model.

## 5 Discussion

In this study, a dynamic graph structure optimization diagnosis (DG-SOD) model was introduced to evaluate its performance on multiple datasets, particularly focusing on the processing of bearing fault datasets. The experimental process is shown in Figure 4. The method not only shows the application potential in academic network analysis, but also shows significant innovation and practical value in the important industrial field of bearing fault diagnosis.



**Fig. 4.** The experimental process

Innovative processing of bearing fault data sets. Bearing is a vital component in industrial machinery, and its health status directly affects the stable operation and production safety of the entire mechanical system. Therefore, the development of effective bearing fault early diagnosis technology is of great significance for preventing equipment failure and reducing maintenance costs. In this context, the proposed DG-SOD model represents a major breakthrough in traditional fault diagnosis methods.

An innovative data processing method is adopted in this study. Firstly, the vibration signal data of the bearing undergoes processing through fast Fourier transform (FFT) [27], which converts it into the frequency domain signal. Subsequently, the k-nearest neighbor algorithm is utilized to construct a graph structure reflecting the complex relationship between the bearing states. The construction method of this graph structure provides a new perspective for the model, so that it can learn deeper and more detailed feature information from the data.

Innovative detection methods are of great significance. The core innovation of the DG-SOD model is that it can effectively capture and utilize the dynamic structural features of graph data, especially the excellent performance in dealing with data imbalance problems and new category detection tasks. On the bearing fault data set, although the accuracy needs to be improved, considering the complexity of the field and the difficulty of fault detection, this result has shown the great application potential of the DG-SOD model.

More importantly, the application of DG-SOD model is not limited to bearing fault diagnosis, and its performance on CiteSeer and Cora datasets further verifies the wide applicability and flexibility of the model. This cross-domain application ability, especially the practical application in the industrial field, indicates that the method is innovative and forward-looking in graph data analysis and processing.

Through the innovative processing of bearing fault data sets and the application of DG-SOD model, the research not only provides an effective novel method for early fault diagnosis of industrial equipment, but also opens up a new way for using graph neural network to process complex graph structure data. The proposal of this innovative detection method is undoubtedly of great significance for promoting technological progress in related fields and realizing more intelligent and automated fault diagnosis systems.

## 6 Conclusion

In this study, the DG-SOD model, a novel graph neural network, is introduced for early fault diagnosis and node classification tasks in bearing fault diagnosis and academic network analysis. By incorporating the k-nearest neighbor algorithm for graph construction and enhancing the Focal Loss function with gDOC-based optimization, DG-SOD effectively addresses class imbalance and new class detection issues. Experiments across various datasets demonstrate DG-SOD's superior performance in classification accuracy, particularly in identifying bearings under different fault conditions and accurately classifying nodes and topics. Future work will focus on optimizing DG-SOD for broader applications, exploring its interpretability, and validating its utility in practical problems like social network analysis and traffic prediction. Through these efforts, DG-SOD is expected to have a significant impact on graph data analysis.

## References

1. Zhang, Y. et al.: G-CoS: GNN-accelerator co-search towards both better accuracy and efficiency. In: 2021 IEEE/ACM International Conference On Computer Aided Design (ICCAD). IEEE (2021).
2. Wang, Y. et al.: Accelerating Distributed GNN Training by Codes. *IEEE Transactions on Parallel and Distributed Systems* 2(5), 99–110 (2023).
3. Ukey, N. et al.: Efficient knn join over dynamic high-dimensional data. In: Australasian Database Conference. Springer International Publishing, Cham (2022).
4. Tian, J. et al.: Sliding Focal Loss for Class Imbalance Classification in Federated XGBoost. In: 2022 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCLOUD/SocialCom/SustainCom). IEEE (2022).
5. Galke, L. et al.: Lifelong learning on evolving graphs under the constraints of imbalanced classes and new classes. *Neural Networks* 164, 156–176 (2023).
6. Hoang, D.-T., Kang, H.-J.: A survey on deep learning based bearing fault diagnosis. *Neurocomputing* 335, 327–335 (2019).
7. Saini, N., Saha, S., Bhattacharyya, P.: Automatic scientific document clustering using self-organized multi-objective differential evolution. *Cognitive Computation* 11, 271–293 (2019).
8. Cai, G. et al.: EMD and GNN-AdaBoost fault diagnosis for urban rail train rolling bearings. *Discrete and Continuous Dynamical Systems-S* 12(4&5), 1471–1487 (2019).

9. Doan, H.-G. et al.: An Efficient Strategy for Catastrophic Forgetting Reduction in Incremental Learning. *Electronics* 12(10), 2265 (2023).
10. Ji, Z. et al.: Coordinating experience replay: A harmonious experience retention approach for continual learning. *Knowledge-Based Systems* 234, 107589 (2021).
11. Fu, S. et al.: Few-Shot Learning With Dynamic Graph Structure Preserving. *IEEE Transactions on Industrial Informatics* 2(5), 99–110 (2023).
12. Alonso-González, M. et al.: Bearing fault diagnosis with envelope analysis and machine learning approaches using CWRU dataset. *IEEE Access* (2023).
13. Kim, S.-W., Gil, J.-M.: Research paper classification systems based on TF-IDF and LDA schemes. *Human-centric Computing and Information Sciences* 9, 1–21 (2019).
14. Gao, Y. et al.: Fault diagnosis of rolling bearings using weighted horizontal visibility graph and graph Fourier transform. *Measurement* 149, 107036 (2020).
15. Fu, S. et al.: Few-Shot Learning With Dynamic Graph Structure Preserving. *IEEE Transactions on Industrial Informatics* 2(5), 99–110 (2023).
16. Lan, S. et al.: An improved iterative update strategy and high-order shape functions for digital image correlation. *Optics and Lasers in Engineering* 176, 108070 (2024).
17. Xue, J., Ma, J.: Extreme Sample Imbalance Classification Model Based on Sample Skewness Self-Adaptation. *Symmetry* 15(5), 1082 (2023).
18. BBidoki, N. H., Mantzaris, A. V., Sukthankar, G.: Exploiting weak ties in incomplete network datasets using simplified graph convolutional neural networks. *Machine Learning and Knowledge Extraction* 2(2), 8 (2020).
19. Amine, O., Mohammed, M.: Graph Oriented Attention Networks. *IEEE Access* (2024).
20. Fu, S. et al.: Example-feature graph convolutional networks for semi-supervised classification. *Neurocomputing* 461, 63–76 (2021).
21. Liu, Y. et al.: Feature fusion using kernel joint approximate diagonalization of eigen-matrices for rolling bearing fault identification. *Journal of Sound and Vibration* 385, 389–401 (2016).
22. Guan, R. et al.: Deep feature-based text clustering and its explanation. *IEEE Transactions on Knowledge and Data Engineering* 34(8), 3669–3680 (2020).
23. Dai, W.: Classification and analysis of literary works based on distribution weighted term frequency-inverse document frequency. In: *Journal of Physics: Conference Series*. IOP Publishing 1941(1), 1–2 (2021).
24. Chen, K. et al.: Turning from TF-IDF to TF-IGM for term weighting in text classification. *Expert Systems with Applications* 66, 245–260 (2016).
25. Li, C. et al.: Meta-learning for few-shot bearing fault diagnosis under complex working conditions. *Neurocomputing* 439, 197–211 (2021).
26. Pan, S. et al.: Joint structure feature exploration and regularization for multi-task graph classification. *IEEE Transactions on Knowledge and Data Engineering* 28(3), 715–728 (2015).
27. McInerny, S. A., Dai, Y.: Basic vibration signal processing for bearing fault detection. *IEEE Transactions on Education* 46(1), 149–156 (2003).