Sensors and Materials, Vol. 35, No. 11 (2023) 3537–3550
MYU Tokyo

S & M 3434

Time-series-based Equipment Failure Diagnosis Mechanism in the Context of Minority Failure Samples

Cheng-Hui Chen,[1] Yung-Kuan Chan,[2*] and Shyr-Shen Yu[1]

[1]Department of Computer Science and Engineering, National Chung Hsing University, Taichung City 407224, Taiwan
[2]Department of Management Information Systems, National Chung Hsing University, Taichung City 407224, Taiwan

(Received July 18, 2023; accepted October 16, 2023)

Keywords: time-series data, equipment failure diagnosis, minority failure samples, hybrid generation, WGAN

Industrial environments frequently encounter complex time-series data such as machine vibration patterns, motor thermal imaging, and sensor pressure metrics. Equipment failure prediction grapples with the temporal nature of the data and the challenge posed by minority failure instances. In this paper, we introduce a refined generative mechanism, building on the foundation of the Wasserstein generative adversarial network (WGAN) and the borderline synthetic minority oversampling technique (Borderline-SMOTE). By utilizing time-series features, the proposed method effectively addresses the intricacies of predictive modeling. To demonstrate its efficacy, we used a complex and multisensor hydraulic system dataset for validation. Experimental results indicate that the proposed method outperforms existing strategies, enhancing the F1 score by at least 2.21% and achieving a recall rate of 95.51%. This suggests a promising direction for enhancing fault prediction in complex industrial settings.

1. Introduction

In manufacturing and production workflows, precise forecasting and identification of machinery malfunctions are vital for preventing unanticipated operational interruptions and maintaining uninterrupted processes. However, the stark minority of failure data frequently results in an imbalance between normal and failure datasets,(1) thereby rendering the development of dependable predictive models considerably challenging. Furthermore, the majority of accessible data comprise complex time-series information,(2) such as machine vibration patterns, motor thermal imaging, and sensor pressure metrics, thereby amplifying the challenges associated with attaining accurate predictions.

In response to these multifaceted challenges, we introduce a novel approach termed the time-series-based equipment failure diagnosis mechanism in the context of minority failure samples (TS-DMMF). The proposed method aims to enhance the accuracy and reliability of equipment failure prediction and diagnosis, despite the constraints of limited and intricate data.

*Corresponding author: e-mail: star90154@gmail.com
https://doi.org/10.18494/SAM4579

ISSN 0914-4935
https://myukk.org/

Over the past few years, the field of imbalanced data research has focused on two primary approaches: imbalanced learning and data generation.[3] The first approach, imbalanced learning, encompasses various techniques designed to address class imbalance. One of the most prominent methods in this category is the synthetic minority oversampling technique (SMOTE),[4] which generates new synthetic samples within the feature space of minority class instances instead of merely duplicating existing minority samples. Several variants of SMOTE have been proposed, including Borderline-SMOTE, wherein a variation generates samples close to the decision boundary. Furthermore, there are numerous other adaptations, such as SMOTE for Nominal and Continuous attributes (SMOTE-NC) for handling mixed data and support vector machine (SVM)-SMOTE for improving synthesis accuracy.[5] Another type of algorithm enhances the importance of the minority class by increasing their weights.[6–8] Lastly, few-shot learning aims to capture the differences between classes from an extremely limited amount of training data, enabling models to effectively recognize and classify samples from new categories.[9] The second approach is data generation, which has gained popularity owing to the rapid development of deep learning research. Data generation, particularly through generative adversarial networks (GANs), creates synthetic minority class samples with similar feature distributions, improving the recognition ability of machine learning algorithms for minority class samples and enhancing classification performance. The generator's goal is to produce seemingly authentic synthetic data, while the discriminator's goal is to distinguish between real and synthetic data. By alternately training the generator and discriminator, the generator learns to create increasingly realistic data.[10–12] In recent research, a hybrid generative model has been developed by combining the advantages of both GANs and SMOTE.[13,14]

Equipment failure data predominantly comprises time-series data; therefore, it is crucial to consider this type of information during data generation to prevent the loss of temporal features. To effectively characterize the cyclical variations in time-series data, an automated approach for feature extraction from cyclical time-series data has been employed, encompassing various aspects such as statistical features, frequency domain features, and shape-based features.[15,16] A parallel feature extraction strategy based on sliding windows has been implemented for time-series feature extraction, facilitating the efficient processing of large-scale time-series data.[17,18] Furthermore, certain studies have integrated data generation with time-series features, thereby enhancing the efficacy of generating minority samples in time-series data.[19]

This paper's primary contribution focuses on designing a time-series-based mechanism for generating minority equipment failure data. The proposed mechanism encompasses two stages. Initially, we amalgamate multiple time-series feature generation methodologies to establish an appropriate time-series feature extraction technique specifically designed for equipment failure data. Subsequently, considering the inherent constraints of Wasserstein GAN (WGAN),[20] which necessitates a substantial volume of equipment failure data for its generator, we incorporate the Borderline-SMOTE approach to synthesize the required data for the generator and employ authentic equipment failure training data for the validation of the discriminator. This strategy effectively addresses the limitations associated with both Borderline-SMOTE (i.e., excessive similarity in the synthesized data) and WGAN (i.e., constraints that require large amounts of data to synthesize). We have compiled the advantages and disadvantages of relevant generative literature in Table 1.

Table 1
Analysis of literature related to data generation.

	Accurate generation of minority data	Diversified generation of minority class data	Mixed-type data	Considered time-series data
WGAN[20]	No	Yes	No	No
Borderline-SMOTE[4]	Yes	No	No	No
SmoteNC-ctGAN[13]	Yes	Yes	Yes	No

The remainder of this paper is organized as follows. In Sect. 2, we introduce the framework for generating minority class data in time series. In Sect. 3, we focus on the experimental design and comparison with similar models. In Sect. 4, we conclude the paper by summarizing its contributions and outlining future research directions.

2. Materials and Methods

In this section, we introduce a mechanism that addresses the issue of equipment failure prediction in the context of imbalanced time-series data. The research framework comprises three parts: first, data collection and time-series data extraction; second, the process of generating minority data; and finally, model prediction. The proposed TS-DMMF framework is shown in Fig. 1.

2.1 Data collection and time-series feature extraction

In this study, we amalgamate methodologies from the literature[15,16,21–23] to execute automated feature extraction on time-series data. This extraction includes statistical features, frequency domain characteristics, and autoregressive features, among others. A sliding-window-based parallel feature extraction strategy is implemented to efficiently process large-scale time-series data. We summarize the commonly used time-series features for equipment failure diagnosis in Table 2.

2.2 Minority data generation process

The scarcity of equipment failure data presents significant challenges in generation. Therefore, we have designed a mechanism for generating minority failure data samples, ensuring both diversity and accuracy. The process of generating minority class data is illustrated in Fig. 2.

The generation process comprises three steps. First, time feature data train the Borderline-SMOTE model, creating boundary samples between normal and abnormal categories. These boundary samples are capable of more effectively differentiating between normal and fault categories. Second, we train the WGAN generator using the fault samples generated by Borderline-SMOTE and the real failure samples. The generator, capable of synthesizing samples from diverse distributions, and the discriminator, which filters the generated samples, ensure the representativeness of the samples. Third, the generated data is merged with the real data, forming a balanced training dataset representing both normal and failure categories.

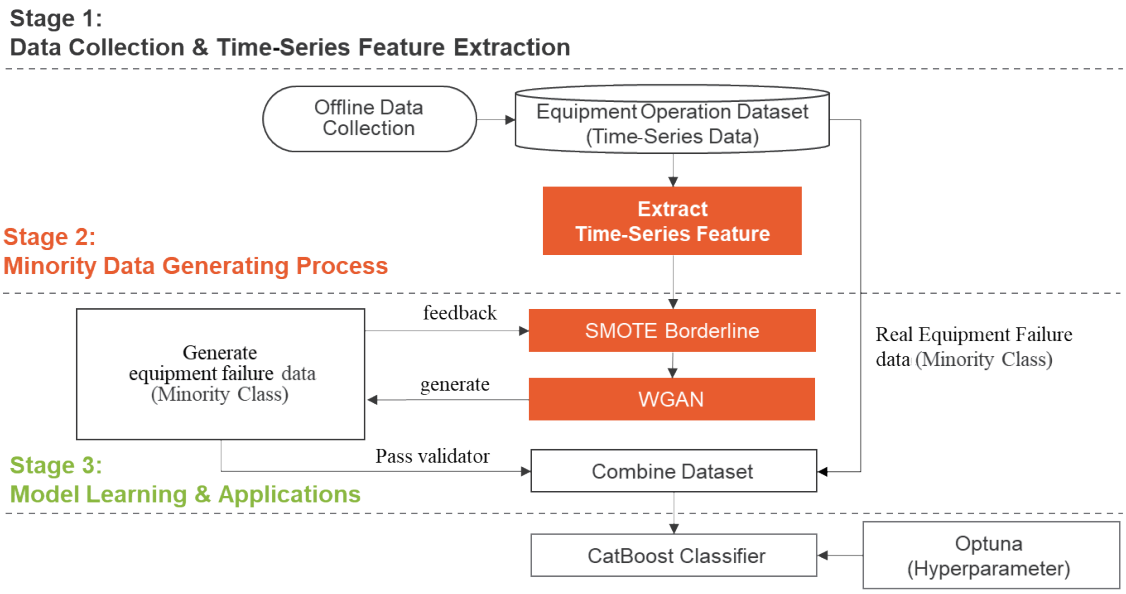

Fig. 1. (Color online) TS-DMMF framework.

Table 2
Analysis of related literature on time-series features.

Category	Feature	Description
Time-domain features	Number of peaks	Evaluates the count of peaks with a minimum support of n within the time series
	Sum of values	Computes the cumulative sum of values across the time series
	Mean	Returns the arithmetic mean value
	Minimum	Identifies the smallest value in the time series
	Maximum	Identifies the largest value in the time series.
	Standard deviation	Reports the standard deviation.
	Skewness	Determines the sample skewness, calculated using the adjusted Fisher–Pearson standardized moment coefficient G1.
	Kurtosis	Determines the kurtosis, calculated using the adjusted Fisher–Pearson standardized moment coefficient G2.
Autoregressive features	Benford correlation	Beneficial for anomaly detection tasks, offering the correlation derived from the first digit distribution when applied
Frequency-domain features	Number of continuous wavelet transform (CWT) peaks	Quantifies the distinct peaks using CWT
	CWT coefficients	Computes the CWT for the Ricker wavelet, also known as the "Mexican hat wavelet."
	Fast Fourier Transform (FFT) aggregated	Presents the spectral centroid (mean), variance, skewness , and kurtosis of the absolute Fourier transform spectrum.

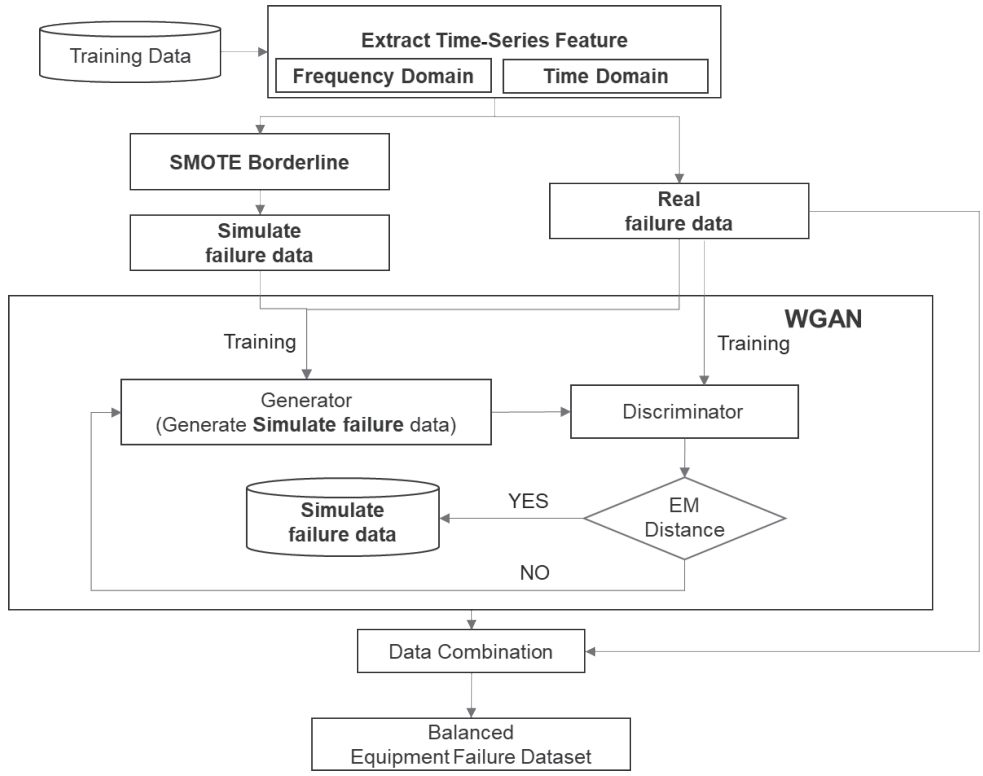

Fig. 2. Process of generating minority class data.

2.2.1 Borderline-SMOTE

In this study, Borderline-SMOTE is utilized.[5] Traditional SMOTE techniques may generate extraneous and ambiguous samples, which can blur the boundaries of classification. In contrast, the Borderline-SMOTE approach selectively synthesizes instances located on the borderline. Consequently, this results in synthesized data that better delineates classes, thereby enhancing the recognition capability for minority classes.[12] The approach is as follows.

Determination of Neighbors for Each Minority Class Sample: A k-NN algorithm is employed, where k is typically set to 5, to identify the k closest neighbors for every minority class sample. Here, z symbolizes the samples within the minority class.

Minority Class Samples Classification into Three Categories (Safe Sample, Boundary Sample, and Noise Sample): For a specific minority class sample z, let m represent the count of majority class samples among its k neighbors. If $m == k$, the sample is categorized as a noise sample. If $0 < m \leq k/2$, it is considered a safe sample. If $k/2 \leq m < k$, it is classified as a boundary sample. The boundary sample set is $X = \{x_1, x_2, ..., x_n\}$.

Synthesis of New Samples: For the boundary sample X, its k neighbors are considered. From these neighbors, a subset of t samples is randomly selected, subject to the constraint $t < k$ and typically set as $t = k/2$. This sample subset, defined as $S = \{s_1, s_2, ..., s_n\}$, facilitates the computation of the novel minority sample X_new. Equation (1) is the calculation formula of X_new.

$$X\_new = x + r * (x - s) \tag{1}$$

Here, the variable definitions are as follows:

- r: a random scalar in the range between 0 and 1,
- x: the designated boundary sample,
- s: the set of neighboring samples corresponding to x.

Borderline SMOTE employs a boundary detection approach for sample synthesis, as illustrated in Algorithm 1.

2.2.2 Wasserstein GAN

The WGAN introduces an innovative solution to lessen the complexities associated with the training of standard GANs and circumvent the prevalent issue of mode collapse.[20] The WGAN adopts a distinctive structure for sample synthesis as demonstrated in Algorithm 2.

The objective function of the discriminator forms the fundamental difference between a conventional GAN and a WGAN. In a conventional GAN, the discriminator is guided by the binary classification of real and synthesized samples. Conversely, the WGAN discriminator

Algorithm 1: Borderline-SMOTE Algorithm (Pseudocode)
Input: Minority class time series feature data z, Number of nearest neighbors k
Output: Generate new minority class samples *ListNew*

for each sample i in minority class samples z:
    Identify the k nearest neighbors for i using k-NN algorithm
for each sample i in minority class samples z:
    Let m be the number of majority class samples among its k neighbors
    if $m == k$:
        i is a noise sample
    else if $0 < m <= k/2$:
        i is a safe sample
    else if $k/2 <= m < k$:
        i is a boundary sample
        for the boundary samples $X = \{x_1, x_2, \ldots, x_n\}$, find their k nearest neighbors and randomly select t samples from k neighbors. This sample subset, defined as $S = \{s_1, s_2, \ldots, s_n\}$
        Generate new minority class samples using Eq(1) where r is a random number between 0 and 1.
        Add *X_new* to *ListNew*
return *ListNew*

Algorithm 2: Wasserstein GAN Algorithm (Pseudocode)
Input: Real fault data $x \sim Dr$, synthetic fault data $y \sim Ds$, discriminator D, generator G, learning rates α_1, α_2
Output: Trained generator G and generated samples *X_new*

Initialize an empty list *X_new* to store generated samples
for each training iteration:
    for each step of discriminator:
        Sample real data x and noise z
        Update D by gradient ascent using x and $G(y)$ to minimize the EM distance using Eq(2)
    Sample noise z
    Update G by gradient descent using $D(G(y))$
    Generate new sample $X\_new = G(y)$
    Add *X_new* to *ListNew*
return *ListNew*

employs the Earth-Mover (EM) distance between real and synthesized sample distributions as its objective function. As a result, the learning task for the WGAN discriminator evolves into a regression problem.[24] The EM distance is defined using Eq. (2).

$$W\left(Dr, Ds\right) = \inf_{\gamma \in \Pi(Dr, Ps)} E_{(x,y)\sim\gamma}\left[\left\|x - y\right\|\right] \tag{2}$$

Here, the variable definitions are as follows:

- The symbol inf denotes the infimum (greatest lower bound).
- The set $\Pi(Dr, Ds)$ denotes the set of all joint distributions $\gamma(x, y)$ whose marginals are Dr and Ds, respectively.
- The symbol $||x - y||$ denotes the L1 norm (Euclidean distance) between x and y.
- The expectation $E_{(x,y)\sim\gamma}\left[\left\|x - y\right\|\right]$ is taken over samples (x, y) drawn according to the joint distribution $\gamma(x, y)$.

The Wasserstein distance represents the minimum mean distance required to shift the mass from one distribution to another, considering the distance as the Euclidean distance between data space points. Often referred to as the EM distance, it can be intuitively comprehended as the minimum expenditure of converting one pile of earth (distribution) to another, where the cost is a function of the quantity of earth moved multiplied by the distance it is moved.

2.2.3 Data combination

In this study, the aim is to reach a balance between fault and normal operation data using synthetic and real fault data. Total data in this research originate from two distinct sources: real data and WGAN generated fault data. The data combination process is illustrated in Algorithm 3.

2.3 Model learning and applications

To validate the efficacy of the proposed method, it is applied to equipment failure prediction and diagnosis. The aim is to measure differences in predictive performance between the unprocessed original dataset and the newly synthesized dataset. There are two types of equipment fault prediction. One is binary classification, where we distinguish between faults and non-faults in the fault diagnosis results. The other is multi-class classification, which classifies diagnosis results into various fault modes based on different faults. As equipment, operation data are often limited in quantity and not ideal for deep learning methods; in this study, we chose the popular CatBoost Classifier method.

Algorithm 3: Data Combination (Pseudocode)
Input: Real failure data *R*, Wasserstein GAN synthesized failure data *W_new*
Output: Balanced failure dataset *T_new*
1. Add *R* and *W_new* to *T_new*
2. **return** *T_new*

2.3.1 CatBoost Classifier and Optuna

CatBoost is an ensemble-learning algorithm that stems from the gradient boosting decision tree.[25] By deploying an ordered boosting strategy and a greedy algorithm, it efficiently addresses issues related to iterative gradient descent, thereby mitigating the risk of overfitting. To enhance model precision and overall performance, we incorporated Optuna for hyperparameter optimization.[26]

2.3.2 Model evaluation

The accuracy of equipment fault prediction is paramount in the evaluation metrics. Conversely, the impact of inaccuracies operations lessens. The confusion matrix is utilized to display the instances of accurate equipment predictions and misclassifications, as defined in the format of Table 3. Simultaneously. The selection of metrics includes recall, accuracy, and F1 score[27] present in Eqs. (3) to (5).

$$\text{Recall rate} = TP/(TP + FN) \tag{3}$$

$$\text{Precision} = TP/(TP + FP) \tag{4}$$

$$\text{F1 Score} = 2/((1/\text{Precision}) + (1/\text{Recall})) \tag{5}$$

3. Results and Discussion

In this section, we outline the experimental design, validation results, and comparisons with similar models.

3.1 Dataset description

To verify the effectiveness of the proposed method in the experiment, we choose a dataset for predicting equipment failure, which contains time-series and imbalanced data.[28] The dataset includes 17 complex sensors, such as pressure, motor power, volume flow, and operating vibration, among others. Moreover, the data is raw, without any feature extraction. The sensors relevant to the hydraulic system and the corresponding fault labels are listed in Tables 4 and 5, respectively.

Table 3
Confusion matrix for fault diagnosis dataset.

		Actual	
		Failure	Normal
Prediction	Failure	*TP* (True positive)	*FP* (False positive)
	Normal	*FN* (False negative)	*TN* (True negative)

Table 4
Sensor description and sampling rate.

Sensor ID	Description	Unit	Sampling rate (Hz)
PS1-PS6	Pipeline pressure	bar	100
EPS1	Motor power	W	100
FS1-FS2	Volume flow	l/min	10
TS1-TS4	Pipeline temperature	℃	1
VS1	Operating vibration	mm/s	1
CE	Cooling efficiency	%	1
CP	Cooling power	%	1
SE	Efficiency factor	%	1

Table 5
Types of valve operational states.

	Status	Number of data
Valve Condition	100: optimal switching behavior	1125 (50.02%)
	90: small lag	360 (16.33%)
	80: severe lag	360 (16.33%)
	73: close to total failure	360 (16.33%)

Valve condition data is imbalanced. The distribution is depicted in Fig. 3, where state 100 signifies normal samples, while other states indicate either deteriorating equipment or equipment fault samples.

3.2 Parameter settings

In this experiment, the parameter settings for WGAN and CatBoost are listed in Tables 6 and 7, respectively.

3.3 Evaluation of generated data

The real and generated data are analyzed by PCA and TSE, as shown in Fig. 4. The pink color represents the real data after PCA and TSE, the gray color represents the synthetic data after PCA and TSE by TS-DMMF, and the deep red color represents regions where the real data and synthetic data generated by TS-DMMF are close or even overlap. The result indicates that the synthetic data generated by TS-DMMF exhibits a similar distribution to the real data.

3.4 Experiment results

In the preliminary stages of model validation, the dataset undergoes meticulous feature engineering, accompanied by rectification of its class imbalance. Following this meticulous preprocessing, the synthesis of minority class samples is initiated. To ensure class equilibrium, the Borderline-SMOTE technique is employed. Subsequent to this balancing procedure, the curated data is fed into a WGAN. Within this architecture, dedicated training models are devised

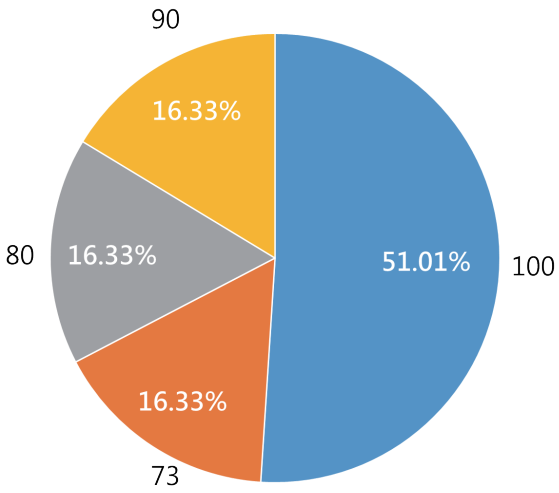

Fig. 3. (Color online) Distribution of valve operational states.

Table 6
Parameter settings for WGAN.[20]

Parameter	Value
Batch size	64
Size of output samples	Generator: (256,256,256) Discriminator: (256,256,256)
Optimizer	Adam
Clipping parameter	0.01
Learning rate	0.00005
Activation	ReLU

Table 7
Parameter settings for CatBoost (parameter optimization based on Optuna).[25]

Parameter	Value
Iterations	72
Learning rate	0.02
Early stopping rounds	10 (Default)
Loss function	MultiClass
Depth	6

explicitly for the minority classes: Class 73, Class 80, and Class 90. Consequently, individual models tailored to each of these three classes emerge. Each model independently produces 765 samples, thereby ensuring a harmonious representation across the dataset.

Data validation was performed using threefold cross-validation. The predicted results of valve condition are shown in Table 8.

The results indicate that the fault prediction results are close to the predicted results for normal operations, indicating that the issue of insufficient fault data has been resolved. Additionally, the overall average prediction accuracy exceeds 95%.

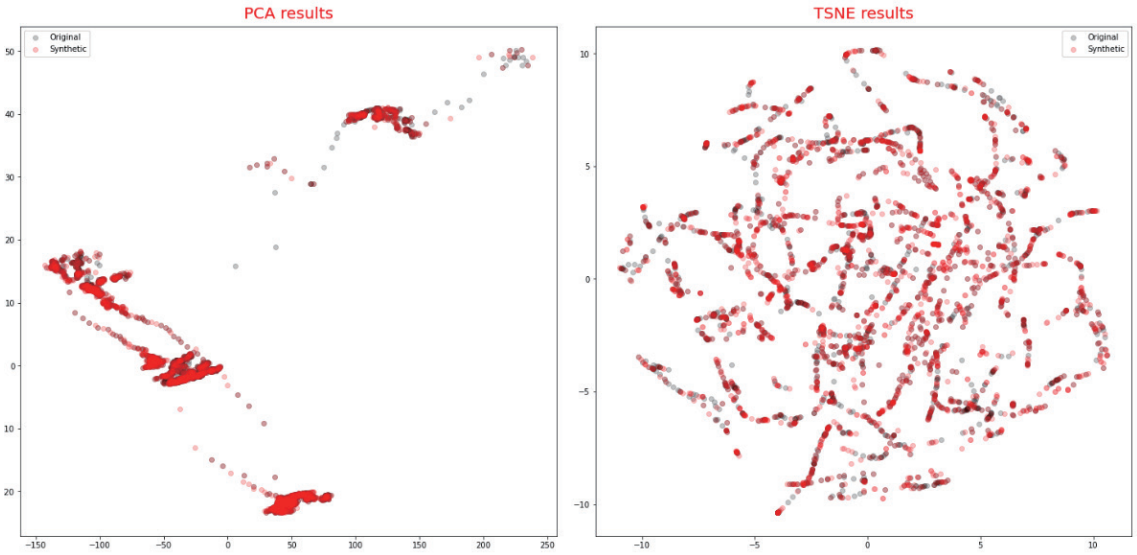

Fig. 4. (Color online) PCA and t-SNE test for real and synthetic data generated by TS-DMMF.

Table 8
Predicted results of valve condition.

Status	Quantity of generated failure data	Precision (%)	Recall (%)	F1 Score (%)
100: optimal switching behavior	0	98.42	98.42	98.42
90: small lag	765	100	90.60	95.07
80: severe lag	765	91.45	96.40	93.86
73: close to total failure	765	92.68	96.61	94.60
Average		95.64	95.51	95.49

Next, we conducted a comparative analysis of different generation methods. First, we generated fault data using various algorithms and subsequently employed the Catboost classifier for prediction. The results are summarized in Table 9.

WGAN achieves the best recall performance with a score of 95.89%, indicating that the generated fault samples are similar to the original ones, helping the classifier to accurately classify fault samples. On the other hand, the Borderline-SMOTE algorithm demonstrates overall stability, but with a lower recall of only 92.59%, which may lead to misclassification of important fault samples and result in significant losses. Finally, the proposed method demonstrated excellent performance with a precision of 95.64% and an F1 score of 95.49%. Notwithstanding a recall rate that is 0.38% less than that of the WGAN, the difference remains within an acceptable margin. Collectively, this strategy displays robust classification efficacy, characterized by a recall of 95.51%—a testament to its precision in fault categorization.

Finally, we juxtapose the proposed method with other time-series processing techniques applied to this dataset. The results are presented in Table 10.

Table 9
Comparative analysis of different generation methods.

	Precision (%)	Recall (%)	F1 Score (%)
WGAN[20]	89.17	95.89	92.11
Borderline-SMOTE[4]	95.23	92.59	93.89
TS-DMMF (proposed method)	95.64	95.51	95.49

Table 10
Comparison of time-series processing methods implemented on this dataset.

	Precision (%)	Recall (%)	F1 Score (%)
Multivariate time-series classification model[18]	92.80	**92.80**	92.80
Bidirectional LSTM[29]	93.63	92.94	93.28
TS-DMMF (Proposed method)	**95.64**	95.51	**95.49**

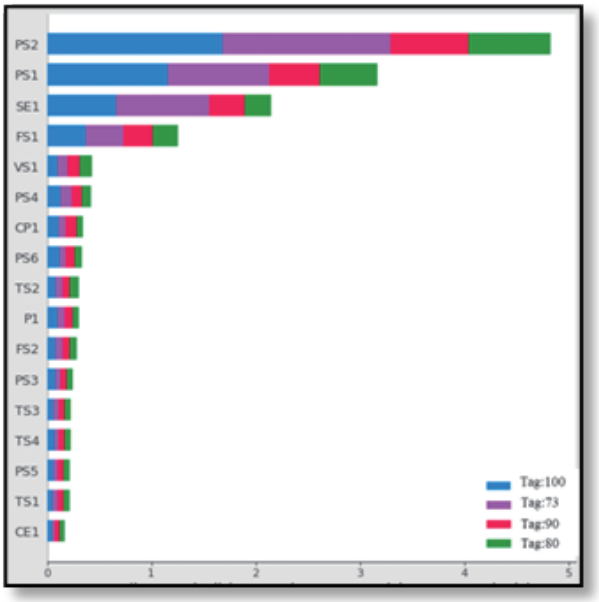

Fig. 5. (Color online) Feature importance analysis in CatBoost classifier.

The purpose of this experiment was to demonstrate the effectiveness of the proposed time-series method. Results indicate that, compared with the time-series method in Bidirectional LSTM,[29] there is an improvement of over 2.21% in the F1 Score. The Recall Rate also reaches 95.51%, suggesting satisfactory fault classification results under time-series data.

3.5 Cause analysis

In the decision-making process of the Catboost classifier, feature importance analysis is employed to pinpoint the critical sensors adept at accurately distinguishing various fault categories. The results are illustrated in Fig. 5.

The top five features are PS2 (Pipeline Pressure), PS1 (Pipeline Pressure), SE1 (Efficiency Factor), FS1 (Volume Flow), and VS1 (Operating Vibration). However, the tags of interest in this study are 73 and 80, as mispredictions of these two could lead to serious unplanned shutdowns. Regrettably, it is not discernible from the figure whether tags 73 and 80 occupy larger portions of the features. The aim of this experiment was to discern feature importance in an attempt to reduce the number of sensors, which aids in computational speed enhancement and cost reduction for data collection.

4. Conclusions

In this paper, we proposed a time-series fault prediction method addressing the dual challenges of time-series data and the imbalance of failure instances. Experimental results indicate that the proposed method outperforms existing strategies, indicating improved fault classification performance, particularly for time-series data. Furthermore, the approach benefits from the inherent explainability of tree-based structures, which simplifies the model interpretation process. This explainability not only aids in identifying the causes of faults but also significantly eases maintenance operations.

References

1 M. Kim, H. Cho, K.-B. Choo, H. Jiafeng, D.-W. Jung, J.-H. Park, J.-H. Lee, S.-K. Jeong, D.-H. Ji, and H.-S. Choi: Sens. Mater. **34** (2022) 3213. https://doi.org/10.18494/sam3991
2 S. K. Singh, S. Kumar, and J. P. Dwivedi: Multimed.Tools Appl. **76** (2017) 18771. https://doi.org/10.1007/s11042-017-4419-1
3 W. Yu, H. Huang, R. Guo, and P. Yang: Sens. Mater. **35** (2023) 2321. https://doi.org/10.18494/sam4509
4 N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer: arXiv [cs.AI]. http://arxiv.org/abs/1106.1813
5 H. M. Nguyen, E. W. Cooper, and K. Kamei: Int. J. Knowl. Eng. Soft Data Paradigms **3** (2011) 4. https://doi.org/10.1504/ijkesdp.2011.039875
6 S. Maliah and G. Shani: Artif. Intell. **292** (2021) 103400. https://doi.org/10.1016/j.artint.2020.103400
7 Y. Geng and X. Luo: arXiv. http://arxiv.org/abs/1801.04396
8 A. Iranmehr, H. Masnadi-Shirazi, and N. Vasconcelos: Neurocomputing **343** (2019) 50. https://doi.org/10.1016/j.neucom.2018.11.099
9 Y. Zhai, Z. Hu, T. Wang, and Y. Wang: Multimed. Tools Appl. https://doi.org/10.1007/s11042-023-15641-1
10 S. Niu, B. Li, X. Wang, and H. Lin: IEEE Trans. Autom. Sci. Eng. **17** (2020) 1. https://doi.org/10.1109/tase.2020.2967415
11 O. Noakoasteen, J. Vijayamohanan, A. Gupta, and C. Christodoulou: IEEE Open J. Antennas Propag. **3** (2022) 488. https://doi.org/10.1109/ojap.2022.3170798
12 R. She and P. Fan: IEEE Internet Things J. **9** (2022) 18589. https://doi.org/10.1109/jiot.2022.3161630
13 C.-H. Chen, C.-K. Tsung, and S.-S. Yu: Appl. Sci. **12** (2022) 9286. https://doi.org/10.3390/app12189286
14 A. Sharma, P. K. Singh, and R. Chandra: IEEE Access **10** (2022) 30655. https://doi.org/10.1109/access.2022.3158977
15 M. Christ, A. W. Kempa-Liehr, and M. Feindt: arXiv [cs.LG]. http://arxiv.org/abs/1610.07717
16 S. Adams, R. Meekins, P. A. Beling, K. Farinholt, N. Brown, S. Polter, and Q. Dong: Prognost. Health Manag. Soc. Conf. **9** (2017). https://doi.org/10.36001/phmconf.2017.v9i1.2452
17 M. Christ, A. W. Kempa-Liehr, and M. Feindt: arXiv. http://arxiv.org/abs/1610.07717
18 A. S. Medishetty, N. S. Muthavarapu, S. G. Goli, B. Sirisha, and B. Sandhya: Lect. Notes Electr. Eng. Springer Singapore. https://doi.org/10.1007/978-981-33-6781-4_7
19 P. Zhao, C. Luo, B. Qiao, L. Wang, S. Rajmohan, Q. Lin, and D. Zhang: Proc. Int. Jt. Conf. Artif. Intell. **3** (2022) 2406. https://doi.org/10.24963/ijcai.2022/334
20 M. Arjovsky, S. Chintala, and L. Bottou: Proc. 34th Int. Conf. Mach. Learn. **70** (2017) 214. PMLR.https://proceedings.mlr.press/v70/arjovsky17a.html

21 T. Schneider, N. Helwig, and A. Schütze: Tm - Tech. Messen **84** (2017) 198. https://doi.org/10.1515/teme-2016-0072
22 A. Kennedy, G. Nash, N. J. Rattenbury, and A. W. Kempa-Liehr: Astron. Comput. **35** (2021) 100460. https://doi.org/10.1016/j.ascom.2021.100460
23 H. Y. Teh, K. I.-K. Wang, and A. W. Kempa-Liehr: IEEE Sens. J. **21** (2021) 18033. https://doi.org/10.1109/jsen.2021.3084970. https://link.springer.com/chapter/10.1007/978-981-33-6781-4_7
24 Q. Wang, X. Zhou, C. Wang, Z. Liu, J. Huang, Y. Zhou, C. Li, H. Zhuang, and J.-Z. Cheng: IEEE Access **7** (2019) 18450. https://doi.org/10.1109/access.2019.2896409
25 A. V. Dorogush, V. Ershov, and A. Gulin: ArXiv 2018. https://doi.org/10.48550/arXiv.1810.11363
26 T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama: Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. https://doi.org/10.48550/arXiv.1907.10902
27 D. M. W. Powers: J. Mach. Learn. Technol. **2** (2011) 37. https://doi.org/10.48550/arXiv.2010.16061
28 N. Helwig, E. Pignanelli, and A. Schutze: 2015 IEEE Int. Instrum. Meas. Technol. Conf. Proc. (2015). https://doi.org/10.1109/I2MTC.2015.7151267
29 K. Kim and J. Jeong: Sensors **20** (2020) 7099. https://doi.org/10.3390/s20247099.

About the Authors

Cheng-Hui Chen is currently pursuing his Ph.D. degree at National Chung Hsing University in Taiwan. He obtained his master's degree in 2012 from National Yunlin University of Science and Technology. Since 2014, he has been serving as a team leader at the Institute for Information Industry. His interests lie in industrial automation, applications of industrial sensors, and data analysis. (star90154@gmail.com)

Yung-Kuan Chan received his Ph.D. degree in computer science from National Chung Cheng University. He is currently a Distinguished Professor at National Chung Hsing University. His research interests lie in image processing, medical imaging technology, and data mining. (ykchan@nchu.edu.tw)

Shyr-Shen Yu received his Ph.D degree from the University of Western Ontario in Canada. He is currently a professor at National Chung Hsing University. His research interests lie in artificial intelligence, image processing, and pattern recognition. (pyu@nchu.edu.tw)