ISSN: 2414-1895 DOI: 10.6919/ICJE.202011_6(11).0014

Prediction of Attenuation State of Marine Propulsion Device Based on Data Drive

Ke Ji[1,a], Xixia Huang[2,b], Zhanbiao Lu, Jiasong Bao

Shanghai Maritime University, ShangHai 201306, China.

[a]jjike21xx@163.com, [b]1227869817@qq.com

Abstract

In order to solve the problem that the current marine propulsion plant attenuation prediction methods need a large number of label data and the sensor data can not directly reflect the state of the components, a ship propulsion plant attenuation and state evaluation method based on principal component analysis ((PCA)) and XGboost prediction algorithm is proposed. Firstly, the measurable parameters which can reflect the performance attenuation of propulsion device are extracted, and the results of PCA feature fusion are taken as predictive health factors to build a model based on XGboost propulsion device attenuation and state evaluation. The experimental results show that the average correct rate of indirect prediction using health factors as labels is 97.5%, slightly higher than the 96.7% of the original label, and higher than the 92.9% correct rate of random forest prediction. It has high precision, which provides a solution to the problem that it is difficult to obtain label data.

Keywords

Data driven; Principal component analysis; XGboost; Ship propulsion device.

1. Introduction

In recent years, with the increase of global trade volume, the volume of maritime transport is increasing rapidly, and shipping has become a very important mode of transportation. Therefore, the prediction and evaluation of the attenuation degree of the ship propulsion device is very important for the safe and efficient operation of the ship and reducing the cost of operation and maintenance [1]. At present, the methods for evaluating the attenuation degree of ship propulsion components are mainly based on physical methods and data-driven methods.

Zheng Weidong [2] through the unified management of fault information, predicting and analyzing the impact of possible faults, he can provide appropriate suggestions for the operation and maintenance of ships. Yu Zhenliang et al. [3] according to the degradation characteristics of mechanical parts, combined with support vector machine and nonlinear Kalman filter to predict the real-time degradation of mechanical parts. Based on the physical method, it is necessary to accurately describe the physical behavior of ship motion to establish an accurate mathematical model.

Peng Xiuyan [4] selects several performance parameters of marine diesel engine as fault features and selects them, then combines KFCM method to determine the fault categories of samples to realize fault diagnosis. V.Manieniyan [5] combines two kinds of neural networks to realize the prediction of engine attenuation state based on ANN model. Niu Xiaoxiao [6] analyzed the factors that may affect the prediction accuracy of neural network, and combined with genetic algorithm to optimize the neural network model for predicting the propulsion performance of marine diesel engine. Zhu Lina [7] uses ReliefF and PCA to filter features to reduce the number of features, and uses SVM to realize fault classification to solve the problem of data imbalance with less actual fault data. Jian-DaWu et

ISSN: 2414-1895 DOI: 10.6919/ICJE.202011_6(11).0014

al.[8] extract the pipeline pressure of engine intake system based on discrete wavelet transform and use generalized regression neural network to realize engine fault diagnosis. FrancescaCipollini[9] discusses the problem of data-driven state-based maintenance. Compared with various advanced supervised learning algorithms, it is proved that supervised learning is effective in predicting the attenuation of ship propulsion devices, but supervised learning requires label data to be deployed.YanghuiTan [10] proposed a mechanical decay state evaluation method based on single-class support vector machine, which only needs a small amount of labeled attenuation data, which reduces the requirement for label data but can not reflect the specific decline. ErikVanem [11] proposed an unsupervised method based on clustering to detect anomalies in diesel engine sensor data, which is also unable to reflect the specific decline state.

Aiming at the problem that it is difficult to obtain the label data of the decay state of the ship propulsion plant, and the unsupervised and semi-supervised algorithms can not reflect the decay state of the components. Starting from the operation data itself, this paper explores the relationship between data characteristics and attenuation coefficient, and adopts the way of indirect prediction to replace the data label [12] with the principal component after feature fusion as a health factor, combined with fuzzy theory. the prediction of the attenuation state of the operation state of the ship propulsion plant are realized.

2. Related technical theory

2.1 Z-Score data standardization

The data set used in this paper includes 16 input parameters and 2 output parameters [13], which come from different components of the propulsion device. When using PCA, we tend to choose some features with large values, and we need to convert the features to the same order of magnitude that can be compared. In this paper, Z-Score is chosen to standardize the data.

2.2 Feature selection

The accuracy of model prediction will be affected by the number of features in the data set. Too many features will increase the running time of the model. It is necessary to select features with high correlation with the results to reduce the number of features when building the model. In this paper, random forest is selected to select the more important features. Using Gini coefficient as a measure of feature importance

2.3 PCA algorithm

PCA algorithm can replace most of the related features with a small number of unrelated new features to realize the feature extraction of data. In this paper, the main information of the data is extracted by PCA algorithm, and the principal component of PCA fusion is taken as the health index of ship propulsion plant.It is generally believed that other principal components can be abandoned when the cumulative contribution rate reaches 85% or more, and that the current principal components can replace the original data

2.4 Principle of XGboost prediction algorithm

XGBoost is a lifting algorithm based on decision tree, which is based on multiple decision trees and is a strong classifier obtained by addition model and forward step-by-step algorithm. Combined with the prediction results of multiple decision trees, XGBOOST improves the prediction accuracy and solves the problem of poor prediction accuracy of a single decision tree.The smaller the value of the objective function, the better. According to the optimization results, the structure of the optimal tree is determined. Finally, the feature with the highest gain is selected as the optimal partition point, and iterated continuously until the gain is less than 0 or the set stop condition is reached

3. General framework

The overall framework of this article is shown in figure 2, which mainly consists of three steps:

ISSN: 2414-1895 DOI: 10.6919/ICJE.202011_6(11).0014

Extraction of health factors of ship propulsion plant. Including data standardization, using PCA algorithm to extract the principal components of ship propulsion unit data.

Construct the prediction model of health factors. It includes verifying the correlation between the fused health factors and the output parameter attenuation coefficients Kmc and Kmt of the data set, and using the XGboost algorithm to establish the prediction model of ship health factors. In the case of unknown data labels, the predicted value of health factors can be obtained by indirect prediction.

Taking the proposed health factors as the evaluation index, the membership function is established, and the fuzzy comprehensive criterion is applied to evaluate the running state of the ship propulsion plant.

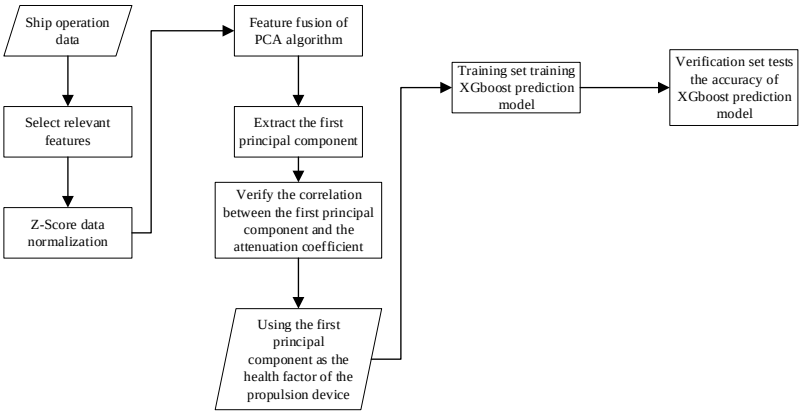

Figure 1. The general framework of attenuation state prediction of ship propulsion unit based on data drive

4. Experimental results and analysis

4.1 Data preprocessing

The data set in this paper is generated by the numerical simulator of naval ships, as shown in Table 1 below. The dataset consists of 11934 records, including 16 input parameters and 2 output attenuation coefficients. Among the input parameters, lp and v, Ts and TP are linearly related, so only lp and Ts features are retained. T1 and P1 are single-valued features, which can be removed directly. The remaining features are selected to delete the two features of the lowest importance. Z-Score normalize the data according to formula (1).

$$z_{ij} = \frac{(x_{ij} - \mu_j)}{\sigma_j} \tag{1}$$

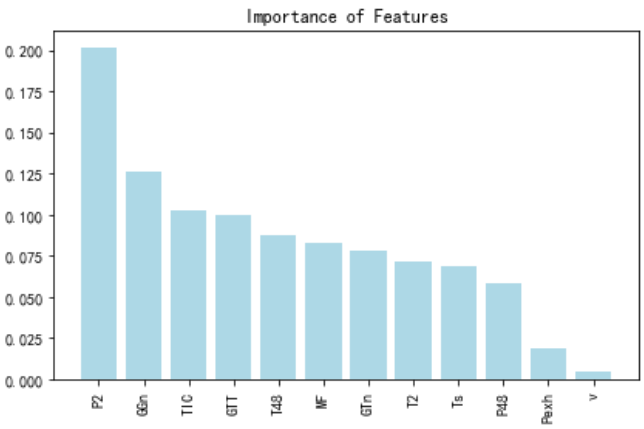

Figure 2 Rank the Importance of Features

ISSN: 2414-1895 DOI: 10.6919/ICJE.202011_6(11).0014

Table 1. Dataset Parameters

Input Parameters		
Parameter	Range	Unit
Lever position	1.138-9.3	/
Ship speed(v)	3-27	kn
LP turbine shaft torque(LPT)	253.547-72784.872	kNm
LP turbine rate of revolutions(LPn)	1307.675-3560.741	rpm
Gas generator rate of revolutions(GGn)	6589.002-9797.103	rpm
Starboard propeller torque(Ts)	5.304-645.249	kNm
Port propeller torque(Tp)	5.304-645.249	kNm
HP turbine exit temoerature(T48)	442.364-1115.797	K
GT compressor inlet air temperature(T1)	288	K
GT compressor outlet air temperature(T2)	540.442-789.094	K
HP turbine exit pressure(P48)	1.093-4.56	bar
GT compressor inlet air temperature(P1)	0.998	bar
GT compressor outlet air temperature(P2)	5.828-23.14	bar
LP turbine exit pressure(Pexh)	1.019-1.052	bar
Combustion chamber injection control(CCIC)	0-92.556	%
Fuel flow(mf)	0.068-1.832	kg/s
Output Parameters		
GT compressor decay state coefficient	0.95-1	/
GT decay state coefficient	0.975-1	/

4.2 Extraction and prediction of health factors

The principal components of ship propulsion equipment data under different speed conditions are extracted by PCA algorithm, and the fused principal components are obtained by setting the number of principal components n_components=3 as health factors. At different speeds, the contribution rate after PCA is shown in Table 2. It can be seen from Table 2 that the contribution rate of each principal component has basically reached 85% or more, and basically contains the information of all the features. PCA1 is the first principal component and PCA2 is the second principal component.

Table 2. Contribution rate of Principal Components

LP	9	12	15	18	21	24	27
PCA1	97.31%	93.48%	93.32%	87.98%	88.93%	88.74%	84.37%
PCA2	2.1%	5.65%	2.36%	10.09%	8.24%	9.09%	8.18%

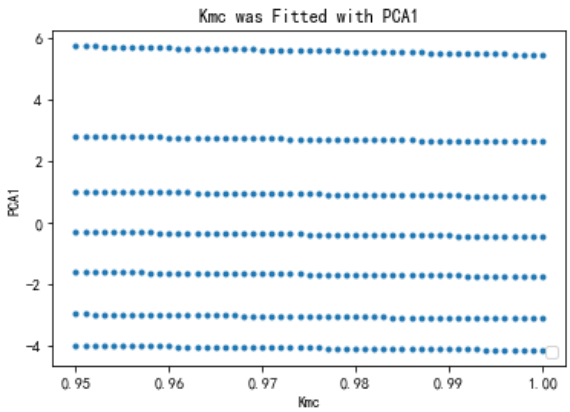

Figure 3. Fitting diagram of attenuation factor and health factor

ISSN: 2414-1895 DOI: 10.6919/ICJE.202011_6(11).0014

4.3 Verification correlation

Verifies that the first principal component as the health factor of the ship propulsion device can characterize the decline of the ship propulsion device and fit the relationship between the first principal component and the ship attenuation factor. The results are shown in figure 3 below. It can be seen that there is a significant linear correlation between the first principal component and the attenuation factor at different speeds. Therefore, the first principal component can be used as a substitute for the attenuation factor as a health factor to characterize the degradation of marine propulsion components.

Table 3: linear correlation coefficient table of health factor and attenuation coefficient KMC.

It can be seen from the above that the health factor proposed in this paper can replace the attenuation factor of the original data set, thus solving the problem that it is difficult to obtain label data in the actual situation. In this paper, the loss rate of marine propulsion plant components is defined, and the formula is as follows

$$E = \left(\frac{PCA\_x - PCA\_\min}{PCA\_\max - PCA\_\min}\right) * 100\% \tag{2}$$

The loss rate of operation data can be obtained from the above formula, which can be used as a label to predict the loss rate of marine propulsion unit components

4.4 Predicted result

Take the velocity is equal to 27 knots as an example, the original data is divided at 7:3, the XGboost algorithm trained in the training set is tested in the test set, and r2_score is used as the model evaluation standard. The prediction accuracy based on the health factors mentioned in this paper is obtained, and the comparison figure is shown in figures 5, 6 and 7

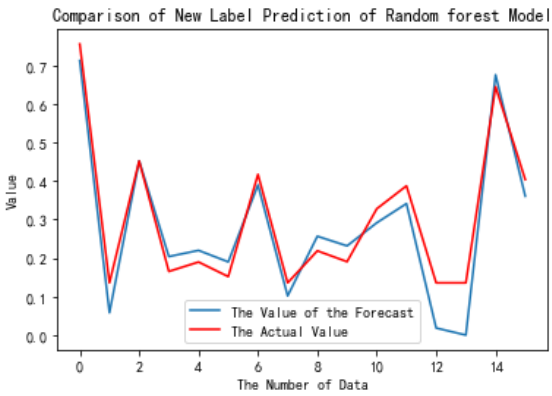

Figure 4. Comparison of New Label Prediction of Random Forest Model

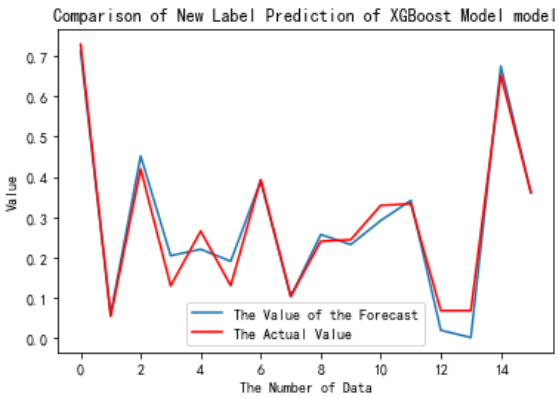

Figure 5. Comparison of New Label Prediction of XGBoost Model

ISSN: 2414-1895 DOI: 10.6919/ICJE.202011_6(11).0014

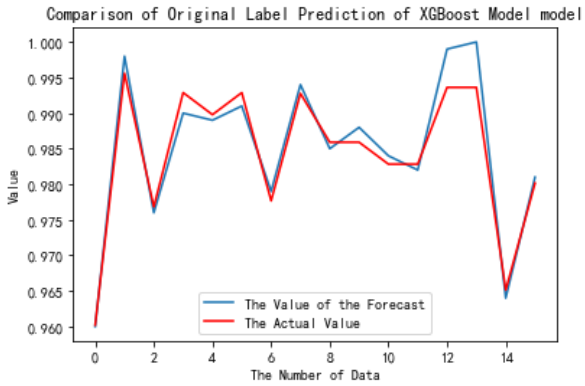

Figure 6. Comparison of Original Label Prediction of XGBoost Model

Table 4.Comparison of the Accuracy of Modeling Methods

Speed(v/knots)	New Label Prediction of Random ForestModel	New Label Prediction of XGBoost Model	Original Label Prediction of XGBoost Model
9	94.0%	97.7%	97.1%
12	93.5%	97.9%	96.8%
15	92.5%	97.1%	97.1%
18	94.5%	97.9%	97.0%
21	91.9%	97.8%	97.2%
24	93.0%	98.2%	97.2%
27	91.4%	96.5%	94.8%

5. Conclusion

The ship propulsion plant is in the state of normal operation most of the time, the fault data samples are very rare, and the label data of the decline state in different stages are even more difficult to obtain. In this paper, through the feature extraction of the operation data of medium and high speed ship propulsion equipment, the feasibility of using principal component as health factor to describe the decline of ship propulsion plant is verified, and the accurate prediction of health factor is realized.The deficiency is that due to the influence of the turbine injection control system (TIC) when running at low speed, the relationship between the principal component and the attenuation factor is difficult to describe, which will be further studied in the following work. To sum up, this paper provides a solution for the attenuation prediction of ship propulsion devices with unknown attenuation label data, and realizes the attenuation prediction of ship propulsion devices at medium and high speed.

Acknowledgments

Natural Science Foundation.

References

[1] Yanghui Tan, Chunyang Niu, Hui Tian,et al. Decay detection of a marine gas turbine with contaminated data based on isolation forest approach[J]. Ship and Offshore Structures, 2020.

[2] Weidong Zheng. Ship equipment Fault Information processing system based on FMEA [J]. Computer and Modernization, 2012 (5): 172-177.

[3] Zhenliang Yu, Zhili Sun, Runan Cao, et al. Research on residual life prediction model of mechanical parts based on support vector machine and Kalman filter [J]. Journal of military Industry, 2018. 39 (5): 991-997.

[4] Xiuyan Peng,Yanyou Chai,Xin Jiang Man. Fault diagnosis of Marine Diesel engine based on PCA-KFCM [J]. Control Engineering, 2012 (19) 2RU 311-315.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202011_6(11).0014

[5] V. Manieniyan, G.Vinodhini, R.Senthilkumar,et al. Wear element analysis using neural networks of a DI diesel engine using biodiesel with exhaust gas recirculation,Energy,Volume 114, 2016, Pages 603-612, ISSN 0360-5442.

[6] Xiaoxiao Niu,Hechun Wang,Xu Li, et al. Optimization of diesel engine performance prediction model based on neural network [J]. Journal of Internal Combustion engine, 2018 Magi 36 (6): 561-568

[7] Lina Zhu,Xuejie Wang,Jinjie Zhang. Research on engine fault diagnosis method based on ReliefF-PCA and SVM [J]. Journal of Beijing University of Chemical Technology (Natural Science Edition), 2018 .45 (1): 55-59.

[8] Wu J-D, Huang C-K, Chang Y-W, et al. Fault diagnosis for internal combustion engines using intake manifold pressure and artificial neural network[J].Expert Systems with Applications. 2010,37(2):949-958.

[9] Francesca Cipollinia, Luca Onetoa, Andrea Coraddu,et al. Condition-Based Maintenance of Naval Propulsion Systems with supervised Data Analysis[J].Ocean Engineering. 149(2018) :268-278.

[10] Yanghui Tan, Chunyang Niu, Hui Tian,et al. A one-class SVM based approach for condition-based maintenance of a naval propulsion plant with limited labeled data[J]. Ocean Engineering. 193 (2019) 106592.

[11] Xiaoqiong Pang, Zhuqing Wang, Jianchao Zeng, et al. Prediction of remaining service life of lithium-ion battery based on PCA-NARX [J]. Journal of Beijing Institute of Technology, 2019. 39 (4): 406-412.

[12] Taoying Li, Miao Hua, and Qian Yin . The Temperature Forecast of Ship Propulsion Devicesfrom Sensor Data[J]. Information, 2019,10,316.