

PAPER • OPEN ACCESS

## Prediction of Bearing Remaining Useful Life based on Mutual Information and Support Vector Regression Model

To cite this article: Wentao Sui *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **533** 012032

View the [article online](#) for updates and enhancements.

# Prediction of Bearing Remaining Useful Life based on Mutual Information and Support Vector Regression Model

Wentao Sui<sup>1</sup>, Dan Zhang<sup>2,a</sup>, Xiaomei Qiu<sup>1</sup>, Wei Zhang<sup>1</sup> and Lin Yuan<sup>1</sup>

1.School of Mechanical Engineering, Shandong University of Technology, 255049 Zibo, China

2.School of Electrical & Electronic Engineering, Shandong University of Technology, 255049 Zibo, China

Corresponding author: zhangdan\_sdut@163.com

**Abstract.** In order to evaluate the degradation state of the mechanical equipment and master the information of the remaining useful life (RUL) of the bearing accurately, this paper presents a method for predicting the remaining useful life of bearings based on mutual information (MI) and support vector regression (SVR) model. The proposed method includes two steps of online and offline, the offline step is used to build a degradation model of the bearing by learning, the online step uses the degradation model to predict the remaining useful life. By analyzing the experimental data of bearing full lifetime degradation, the results show that the method can effectively simulate the bearing degradation process and predict the remaining useful life of the bearing.

## 1. Introduction

Prognostics and health management (PHM) of industrial systems is the core activities of intelligent maintenance such as condition based maintenance (CBM) and predictive maintenance (PM). It includes state monitoring, fault detection, fault diagnosis, fault prediction and decision support. Therefore, fault prediction and health management of components can improve machine availability, reliability and safety. The purpose of fault prediction and health management of rotating machinery is not only to detect faults, but also to predict the time when the machine can run safely and realize its functions.

At present, many literatures are focused on bearing fault diagnosis [1-3], the study of remaining useful life prediction has gradually attracted the attention of scholars [4]. Kang [5] use a combination of multiple evaluation criteria to analyze the validity of features, select the effective features accurately and comprehensively. Shen [6] proposed a new method of remaining useful life prediction based on relative characteristics by using the relative root mean square value which is not affected by bearing individual differences to evaluate the bearing performance degradation law. Lei [7] took wavelet packet decomposition technology to extract the frequency band energy ratio of vibration signal as the sensitive index of bearing degradation. The running trend of bearing was predicted by choosing the radial basis function and appropriate parameters. In the above methods the calculation complexity is increased by using multiple evaluation criteria to select features, while the bearing fault information contained in a single indicator is too little to accurately predict the bearing remaining useful life. Therefore, this paper uses mutual information method [8] to find a few effective features which can



represent a large number of original features, it provides more remaining useful life information, and then predict the bearing remaining useful life better.

To solve the above mentioned problems, this paper proposes a method of mutual information (MI) for nonlinear feature reduction, and combines nonlinear support vector regression to construct health indicators to evaluate the health status of bearings to predict the remaining useful life.

## 2. Feature extraction and reduction

### 2.1. Feature Extraction And Processing

One of the main difficulties in predicting the remaining useful life of bearing is how to establish the characteristic indicators representing the degradation state from the effective feature. In order to accurately predict the bearing remaining useful life, this paper selects several characteristic indicators from time domain and frequency domain as the input sample of the model. These characteristic indicators reflect the running trend and development condition of bearing to some extent, and select the fault sensitive feature to predict the bearing life, as shown in Table 1.

**Table 1.** The features description.

Feature	Feature name
<b>Time domain</b>	peak value、variance、root mean square、skewness index、kurtosis index、 waveform index、peak value index、pulse index、margin index
<b>Frequency domain</b>	mean frequency、frequency center、Standard deviation frequency、Root mean square frequency、kurtosis frequency

Feature processing includes two parts.

(1) The original vibration signals of the bearing contain more noise, the curves drawn by the extracted features will have many "burrs", deal with the sliding average of the features to in order to obtain a relatively smooth curve and reduce noise. 7 points slip average is shown in formula (1).

$$x_n^{MA} = \begin{cases} \frac{1}{n+3} (x_1 + \dots + x_{n+3}) & n \leq 3 \\ \frac{1}{7} (x_{n-3} + \dots + x_{n+3}) & 4 \leq n \leq N-3 \\ \frac{1}{N-n+4} (x_{n-3} + \dots + x_N) & N-2 \leq n \leq N \end{cases} \quad (1)$$

Where  $x_n$  is the original characteristic sequence,  $x_n^{MA}$  is the new sequence after the slip average,  $N$  is the total number of vibration data.

(2) In order to reduce the influence of large differences of characteristic variables on the performance of the model, the extracted feature is normalized.

$$X_{norm} = (Y_{max} - Y_{min})(X - X_{min}) / (X_{max} - X_{min}) + Y_{min} \quad (2)$$

Where  $X_{norm}$  is the result of normalization,  $Y_{max} = 1$ ,  $Y_{min} = 0$ ,  $X$  is the eigenvalue,  $X_{max}$  is the maximum feature value,  $X_{min}$  is the minimum feature value.

### 2.2. Feature Reduction

Mutual information is a basic concept in information theory and a measure of statistical correlation between two random variables. It is usually used to describe the statistical correlation between two systems, or the amount of information contained in another system, as a measure of the amount of

information that systems provide to each other. Mutual information is used to measure the degree of correlation between two variables, representing the part with the same information between two variables. Give the random variables  $X$  and  $Y$ , their marginal probability distribution and joint probability distribution are respectively  $p(x_i)$ ,  $p(y_j)$  and  $p(x_i, y_j)$ , the mutual information is defined as:

$$I(X;Y) = E[I(x_i, y_j)] = \sum_{x_i \in X} \sum_{y_j \in Y} p(x_i, y_j) \log \frac{p(x_i, y_j)}{p(x_i)p(y_j)} \quad (3)$$

Where  $p(x_i)$  is the occurrence probability of  $x_i$ . When  $X$  and  $Y$  are irrelevant variable or independent variables, the  $I(X;Y)=0$ , it means that there is no information contained between the two variables; conversely, the higher the degree of correlation between the two variables, the bigger the mutual information value, the more the same information is included.

In this paper, the mutual information between each input sample feature and the bearing remaining useful life are calculated, preset a threshold. When the mutual information value and the correlation coefficient absolute value are greater than the threshold value, the feature is retained, otherwise it is eliminated.

### 3. SVR prediction theory

Support vector machine is a learning algorithm which is first proposed by Vapnik to deal with small samples. Support vector regression has been successfully applied to various machine learning problems, especially regression problems, sunspot frequency prediction and drug discovery. In this paper, support vector regression is used to study the nonlinear degradation model of bearings. The objective of support vector regression is to estimate the relationship between input and output random variables on the assumption that the joint distribution of input and output variables is completely unknown.

Support vector machines are mainly divided into two types: support vector classification (SVC) and support vector regression (SVR), SVR is the most common application of support vector machines. The model created by SVR technology only depends on a subset of the training data because the cost function constructed by the model ignores all training data close to the prediction threshold of the model. Regression estimation can be formally transformed into a inference problem of the function

$y=f(x)$ , the training set  $X = \{(x_i, d_i), i=1, \dots, l\}$  is given,  $x_i \in R_n$  is the input variable,  $d_i \in R$  is the predictive value and  $l$  is the number of the training set. The regression function is

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) k(x_i, x) + b \quad (4)$$

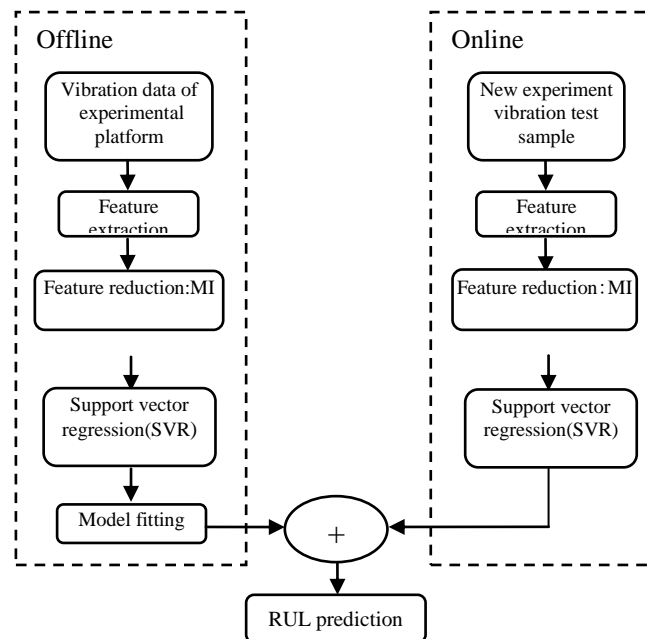
Where  $k(x_i, x)$  is the kernel function,  $b$  is the bias threshold,  $\alpha_i$  and  $\alpha_i^*$  are the Lagrangian multiplier,  $(\alpha_i - \alpha_i^*)$  can get by solving the following optimal problem.

$$\begin{aligned} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)k(x_i, x_j) + \varepsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) - d \sum_{i=1}^l (\alpha_i - \alpha_i^*) \\ s.t. \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \quad \alpha_i, \alpha_i^* \in [0, C] \end{aligned} \quad (5)$$

Where  $\varepsilon$  is insensitive factor,  $d$  is the order of kernel function and  $C$  is penalty factor.

### 4. Prediction process

Prognostics and health management is the core activity of condition based maintenance and predictive maintenance. Predicting remaining useful life can be done by three main methods: model-based, data-driven and hybrid prognostics.



**Figure 1.** Flowchart of remaining useful life prediction.

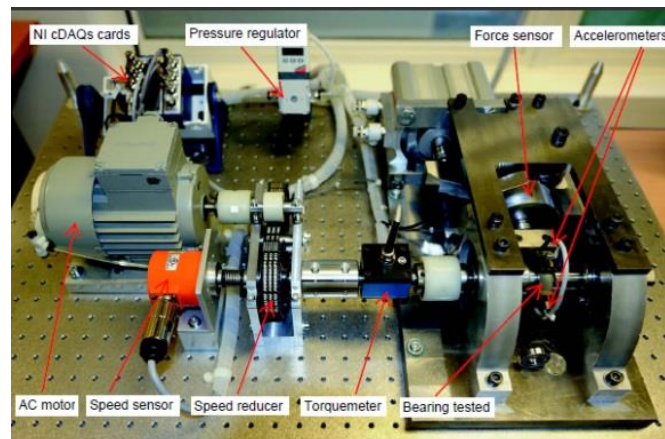
The model-based prediction results are more accurate, but they are more difficult to implement because physical modeling is not a simple task in most applications. Data-driven fault prediction relies on the data provided by the sensor to extract features, which are used to build remaining active life models. This method is easy to implement, but the prediction results are not as accurate as the former, which provides a compromise between accuracy and complexity. The hybrid method combines the advantages of the first two methods, but also has these disadvantages.

The method presented in this paper belongs to data-driven, the framework of life prediction is shown in figure 1. In the off-line step, the extracted features of bearing vibration signal are reduced by mutual information method. The combined features are taken as the input of SVR model, and the ratio of current remaining useful life of bearing to full life is model output.

## 5. Experimental verification

### 5.1. Experiment platform

PRONOSTIA experimental platform is used to test and verify bearing health assessment, fault diagnosis and fault prediction models, as shown in figure 2. The experimental platform can accelerate the bearing degradation in only a few hours and collect the real experimental data describing the bearing degradation law. The design of the PRONOSTIA platform can provide the bearing degradation data under various operating conditions [9].



**Figure 2.** Experimental platform

The degradation of bearing is mainly realized by two kinds of sensors installed on the experimental platform: acceleration sensor and temperature sensor. The inner ring of the bearing rotates with the shaft, the outer ring is fixed, and the acceleration sensor and temperature sensor are mounted on the outer ring of the bearing. The two accelerometers are installed in the horizontal and vertical position of the outer ring of the bearing, and the vibration information is collected from the horizontal direction and the vertical direction respectively. The sampling frequency of acceleration sensor and temperature sensor is 25.6 kHz and 0.1 Hz respectively.

On this platform, the bearing life acceleration experiments are carried out under three different working conditions, and the rotational speed and load information of each condition is shown in Table 2.

**Table 2.** Working condition information.

Condition	Speed(RPM)	Loading(N)
<b>Condition 1</b>	1800	4000
<b>Condition 2</b>	1650	4200
<b>Condition 3</b>	1500	5000

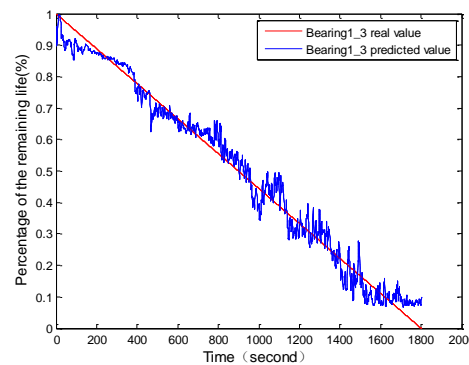
### 5.2. The results and analysis

In this paper, mutual information method is used to reduce the features extracted and the correlation coefficient method (RC, Relation Coefficient) [10] is used to compare and analyze with the former. The two methods are used to calculate the mutual information value and correlation coefficient between each input sample feature and the bearing remaining useful life respectively. A threshold value is preset. When the absolute value of mutual information value and correlation coefficient is greater than the threshold value, the feature is retained, otherwise it is eliminated. With a threshold of 0.89, the selected features and predictive indicators of the two methods are shown in Table 3.

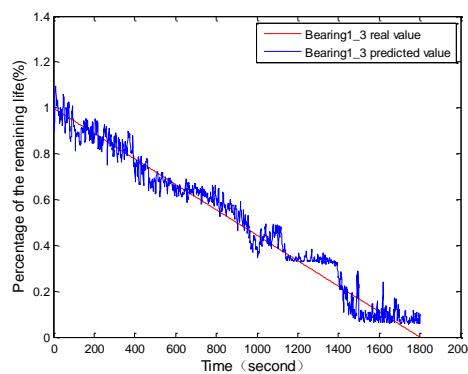
**Table 3.** Feature and prediction result indicators value

method	feature	MI or RC	MSE	R
<b>MI</b>	Margin indicator	0.9002	0.0115	0.9267
	Standard deviation	0.9449		
	frequency	0.9522		
	Kurtosis frequency			
<b>RC</b>	Standard deviation	-0.8963	0.0357	0.8977
	frequency	-0.8934		
	Root mean square	-0.8971		
	frequency			
	Kurtosis frequency			

In this paper, the full life data of bearing 1\_1 in working condition 1 is used as the training set to establish a prediction model to predict the remaining useful life of bearing 1\_3. The sensitive feature combination after bearing reduction is used as the input of the SVR training model and the ratio of remaining useful life to full life is used as the output at present. The prediction results of remaining useful life based on two feature reduction methods are shown in fig.3 and fig.4 respectively. The horizontal coordinate is the acquisition time and the vertical coordinate is the percentage of the remaining useful life.



**Figure 3.** RUL prediction result based on MI



**Figure 4.** RUL prediction result based on RC

It can be seen from the above two diagrams that there is a certain error between the prediction result and the actual value of the bearing remaining useful life, the predicted value fluctuates up and down near the actual value, and the overall trend of the prediction is consistent with the actual value. From Table 3, the squared correlation coefficient ( $R$ ) of the prediction result in the model established by the feature reduction of MI method is 0.9267, which is higher than the  $R$  value of the model established by the feature reduction of the RC method, the mean square error (MSE) of the model is 0.0115, less than that of the latter. The advantages of the feature reduction method proposed in this paper in predicting the remaining useful life are proved.

## 6. Conclusion

Aiming at the difficulty of predicting the bearing remaining useful life, the feature reduction method of bearing signal is improved, a prediction model of remaining useful life based on mutual information and support vector regression is proposed. Firstly, reduce the extracted bearing signal features, select several effective features by mutual information method, which form the support vector regression

model of sample set with remaining useful life, then the model is used to predict the remaining useful life of other bearings.

### Acknowledgments

This research was supported by Shandong Provincial Natural Science Foundation (China, approved grant number ZR2016EEM20, ZR2016FL15).

### References

- [1] W.T. Sui, D. Zhang, W.Wang. Fault Diagnosis of Rolling Element Bearings based on EMD and MKD [J]. Journal of Vibration and Shock, 2015, 34(9): 55-59+64
- [2] W.P. Liu, Y.Q. Liu, S.P. Yang, X.H. Gu. Fault diagnosis of rolling bearing based on typical correlated kurtogram[J].Journal of Vibration and Shock,2018, 37(8):87-92
- [3] D. Zhang, W.T. Sui, Q.J. Guo. Condition Monitoring of Rolling Element Bearings Using Optimal Gabor Filters[J]. Journal of Beijing University of Posts and Telecommunications, 2016: 39(2):103-107
- [4] T Benkedjouh, K Medjaher, N Zerhouni, S Rechak. Remaining useful life estimation based on nonlinear feature reduction and support vector regression[J]. Engineering Applications of Artificial Intelligence,2013, 26: 1751–1760
- [5] Sh.Q. Kang, L.Q. Ye, Y.J Wang, J.B. Xie, Mikulovich V I. Remaining useful life prediction of rolling bearing based on MCEA-KPCA and combined SVR[J]. Journal of Electronic Measurement and Instrumentation,2017,31(9): 1365-1371
- [6] Zh.J. Shen, X.F. Chen, Zh.J. He, Ch. Sun, X.L. Zhang, Zh.W. Liu. Remaining useful life Predictions of Rolling Bearing Based on Relative Features and Multivariable Support Vector Machine[J].Journal of Mechanical Engineering, 2013, 49(2): 183-189
- [7] Y.G. Lei, X.H. Wu, Ch. Wu, J. Lin. Research On Trend Prediction of Rolling Element Bearings Based on Relevance Vector Machine[C]// National Conference on Equipment Monitoring, Diagnosis and Maintenance, National Equipment Failure Diagnostic Conference and 2014 National Equipment Diagnostic Engineering Conference. 2014
- [8] J. Li, J. Li, Sh. Y. Zhang, H.B. Shu. Mechanical Fault Method Using EEMD Algorithm and Mutual Information Method[J].Journal of Huaqiao University(Natural Science), 2018, 39(1): 7-13
- [9] Nectoux P, GOURIVEAU, Medjaher K ,et al. Pronostia: An experiment platform for bearings accelerated degradation tests[C].Proceedings of IEEE International Conference on Prognostics and Health Management, 2012: 1-8
- [10] Y.F. Zhen, H.J. Shi. Attribute reduction algorithm based on relation coefficient and conditional information entropy[J]. Computer Engineering and Applications, 2011, 47(16):26-28