

# Similarity-based Feature Extraction from Vibration Data for Prognostics

Huang, Wei<sup>1</sup>, Farahat, Ahmed<sup>2</sup>, and Gupta, Chetan<sup>3</sup>

<sup>1,2,3</sup> *Industrial AI Lab, Hitachi America, Ltd. R&D, Santa Clara, CA, 95054, U.S.A*  
*First.Last@hal.hitachi.com*

## ABSTRACT

Many techniques for prognostics depend on estimating then forecasting health indicators that reflect the overall health or performance of an asset. For vibration data, health indicators are typically calculated by combining various vibration measurements along with derived features extracted from time, frequency or time-frequency domain analysis. However, selecting or handcrafting good features is a labor-intensive task. On the other hand, deep learning models might be able to learn health indicators automatically from vibration data but require large amount of training data, which are typically hard to obtain from real assets. In this paper, we propose an innovative similarity-based feature extraction method for vibration data which can then be used to learn health indicators and estimate remaining useful life of equipment. The method learns a set of representative templates of frequency spectra for both normal and failure states, and then calculates similarity-based features between new vibration data and the set of learned templates. These features are used to estimate health indicators which are then extrapolated to estimate the future health condition of the asset and its remaining useful life. The proposed method has been tested on the PRONOS-TIA bearing dataset provided by FEMTO-ST Institute and achieved a higher accuracy in estimating the remaining useful life of bearings compared to other studies. The results demonstrate the effectiveness of the proposed method for assets with limited training data.

## 1. INTRODUCTION

Prognostics is concerned with predicting the future health or performance of an asset and estimating its remaining useful life. This task is referred to as Remaining Useful Life (RUL) estimation. Accurate prediction of RUL prevents unexpected failures, eliminate costly repairs, and accordingly increase asset availability.

Vibration measurement has been widely studied for prognostics tasks for industrial equipment, especially rotating machinery and equipment with bearings, motor and gearbox, base on time, frequency or time-frequency domain analysis (Plante et al., 2015; Doguer & Strackeljan, 2009; Singhal & Khandekar, 2013; Al-Badour et al., 2011). Some implementations aim to develop Health Indicators (HIs) from vibration data that represent the trends related to degradation condition from normal to failure based on domain expert knowledge or feature extraction methods. In (Doguer & Strackeljan, 2009), time domain features extracted from vibration measurements are used to detect the roller bearing surface defects. Frequency domain analysis performs effectively for stationary signals. In (Plante et al., 2015), by comparing the frequency spectrum of the vibration data for three fault cases (unbalance, mechanical looseness and bearing fault) to that of a healthy motor, the specific natural frequency corresponding with each fault condition are identified. In (Al-Badour et al., 2011), wavelet analysis is investigated for feature extraction from rotating machinery with non-stationary vibration measurements.

However, different features work for different problems. Given a new problem, constructing health indicator by selecting good features from all extracted features might require laborious preliminary analysis. Handcrafting health indicator or features requires domain knowledge which might be hard to get in real application. On the other hand, deep learning is gaining popularity due to the robustness and superiority in terms of accuracy when trained with huge amount of data. Using deep learning method, health indicator can be automatically constructed from vibration data. In the work of (Zhang et al., 2020), implementations of various deep learning algorithms on vibration data for bearing fault diagnostics were reviewed, including Convolution Neural Network (CNN), Deep Belief Network (DBN), and Recurrent Neural Network (RNN). However, deep learning methods require large amount of training data, which are typically hard to obtain from real assets.

In this paper, we propose an innovative similarity-based fea-

---

Wei Huang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ture extraction method for vibration data which can then be used to learn health indicators for prognostics. The method uses Fast Fourier Transform (FFT) to learn a set of representative templates of frequency spectra for both normal and failure states. Then, for a new snapshot of vibration data, similarity measures are calculated based on correlation coefficient and Euclidean distance between the new snapshot and the set of learned templates. These similarity-based features reflect the deviation of the new vibration snapshot from vibration in normal and failure conditions. These features are used to estimate health indicators using a model of multiple Long Short Term Memory (LSTM) networks. The health indicators are then extrapolated to estimate the future health condition of the asset and the remaining useful life. The proposed method is explained in Section 2.

The proposed method is then tested on the PRONOSTIA bearing dataset provided by FEMTO-ST Institute for RUL estimation (Nectoux et al., 2012). The PRONOSTIA bearing dataset is a popular benchmark dataset for RUL estimation since its usage in PHM 2012 data challenge. The winner in the PHM 2012 data challenge presents three methods with different features extracted including spectral kurtosis, various time-frequency domain features using wavelet transform, and human defined features after thorough inspection of the data (Sutrisno et al., 2012). Another health indicator named weighted minimum quantization error crafted by fusing multiple features is used in a model-based RUL prediction method with Particle Filter based algorithm (Lei et al., 2016). In (Guo et al., 2017), related-similarity features are extracted to construct health indicators using RNN but only the similarities to the normal state are considered. In (Chen et al., 2019), the frequency spectrum of the vibration measurements is used directly as input to a RNN based encoder–decoder framework to estimate health indicator and then predict RUL. In the proposed method, similarity-based features from frequency domain are extracted by considering both the normal and failure state. Meanwhile, various typical time domain features are also included. Since the learning data in PRONOSTIA dataset is limited with only 2 run-to-failure bearings for each of the three operating conditions, we construct an ensemble model which contains multiple LSTM networks to estimate the health indicator from the extracted features. The RUL prediction result for the PRONOSTIA bearing dataset produced by our method has a higher accuracy compared to the state-of-the-art studies. The case study on the PRONOSTIA bearing dataset is described in Section 3.

## 2. METHOD

In this paper, we propose an innovative similarity-based feature extraction method for vibration data which can then be used to learn health indicators and estimate the remaining useful life. This section is organized as follows. Section 2.1 introduces the similarity-based feature extraction method,

which serves as data preprocessing for the followed prognostic model. Section 2.2 describes the prognostic model with LSTM networks for health indicator estimation. In model testing stage, the trained HI estimation model is applied to estimate health indicator for any snapshot of vibration data from a new device. The estimated health indicators are then used to predict remaining useful life for the new device, which is described in Section 2.3

### 2.1. Feature Extraction

In the proposed feature extraction method, we first learn a set of representative templates of frequency spectra for both normal and failure states using Fast Fourier Transform (FFT). Then, for a new vibration snapshot, similarity measures are calculated based on correlation coefficient and Euclidean distance between the spectra of the new snapshot and the learned templates. These similarity measures serve as similarity-based features which reflect the deviation of the new vibration snapshot from vibrations in normal and failure conditions.

Usually, vibration data is measured continuously at a predefined sampling frequency from an asset and processed as a sequence of snapshots, where each snapshot corresponds to the vibration data measured in a predefined time window. The sampling frequency for vibration measurements could vary in an extremely large span from Hertz level up to MegaHertz level. Higher sampling frequency could cover frequency information in a wider frequency range which allows for more information for building prognostic model. Typically, the sampling frequency is within a few hundred Hertz to a few hundred KiloHertz. For the cases with relatively large sampling frequency, using the original frequency spectrum directly as features results in large dimensionality in feature space and may degrade the performance of prognostic models. Therefore, we consider the similarity measures on top of the original frequency spectrum as features which carry the information of deviation from normal state or approach to failure state.

Given continuous vibration measurements from a run-to-failure asset, the initial period is usually considered as normal condition and the period approaching failure is considered as failure condition. In extracting similarity-based features from frequency spectrum, the representative spectrum templates for normal and failure condition are first learned by averaging spectra from a number of vibration snapshots measured in normal condition and failure condition respectively. The spectrum templates for normal condition  $\mathbf{F}^n$  and failure condition  $\mathbf{F}^f$  are two series represented as:  $\mathbf{F}^n = [\mathbf{f}_1^n, \dots, \mathbf{f}_i^n, \dots, \mathbf{f}_m^n]$  and  $\mathbf{F}^f = [\mathbf{f}_1^f, \dots, \mathbf{f}_i^f, \dots, \mathbf{f}_m^f]$ , which span the frequency range below half of the sampling frequency based on Nyquist theorem.

For a new vibration snapshot, the similarities are measured between its frequency spectrum  $\mathbf{F} = [\mathbf{f}_1, \dots, \mathbf{f}_i, \dots, \mathbf{f}_m]$  and

the spectrum templates for normal ( $\mathbf{F}^n$ ) and failure ( $\mathbf{F}^f$ ) condition respectively. Two type of similarities are estimated including Pearson correlation coefficient and Euclidean distance. Pearson correlation coefficient is a normalized statistic measure for the linear correlation between two series as shown in Eq. 1. The Euclidean distance is a popular measure for the point-wise distance between two series without considering any lag or distortion as shown in Eq. 2..

$$\rho(\mathbf{X}, \mathbf{Y}) = \frac{\text{cov}(\mathbf{X}, \mathbf{Y})}{\sigma_{\mathbf{X}}\sigma_{\mathbf{Y}}} \quad (1)$$

$$d(\mathbf{X}, \mathbf{Y}) = \sqrt{\frac{1}{n} \sum_{i=1}^n (\mathbf{X}_i - \mathbf{Y}_i)^2} \quad (2)$$

Given the case with vibration data measured in a high sampling frequency, the full spectrum is equally divided into a number of sub-frequency bands and similarities are calculated for each of the sub-bands.

Besides the similarity-based features, eight typical statistical time domain features are also included for prognostics. Given a snapshot of 1-dimension raw vibration data:  $\mathbf{v} = [\mathbf{v}_1, \dots, \mathbf{v}_j, \dots, \mathbf{v}_n]$ , the eight statistical time domain features are calculated from the absolute value in  $\mathbf{v}$  as defined in Table 1.

Table 1. Statistical Time Domain Features

Feature	Definition
Mean( $ \mathbf{v} $ )	$\frac{1}{n} \sum_{j=1}^n  v_j $
Min-to-Max	$\max( \mathbf{v} ) - \min( \mathbf{v} )$
Standard Deviation ( $SD$ )	$\sqrt{\frac{\sum_{j=1}^n ( v_j  - \overline{ v })^2}{n}}$
Skewness	$E[(\frac{ v  - \overline{ v }}{SD})^3]$
Kurtosis	$E[(\frac{ v  - \overline{ v }}{SD})^4]$
Energy	$\sum_{j=1}^n  v_j ^2$
Shannon entropy ( $H$ )	$H( \mathbf{v} ) = -\sum_{i=1}^m \mathbf{p}_i \log_2(\mathbf{p}_i)$
Energy/Entropy	$Energy/H$

where  $E$  is the expectation operator

## 2.2. Health Indicator Estimation using LSTM Networks

After feature extraction, we estimate health indicator from features using LSTM networks. We assume that the training dataset includes a set of devices  $[\mathbf{D}_1, \dots, \mathbf{D}_i, \dots, \mathbf{D}_m]$ , where each device has one run-to-failure sequence of features. In training the HI estimation model with LSTM networks, the training data sequence from all devices are formatted to be a set of data samples:  $\mathbf{X}, \mathbf{Y} = \{\mathbf{X}_t^i, y_t^i\}$ , ( $t \in [0, \dots, T^i], i \in [1, \dots, m]$ ), where  $t$  is the time step and  $T^i$  denotes the failure

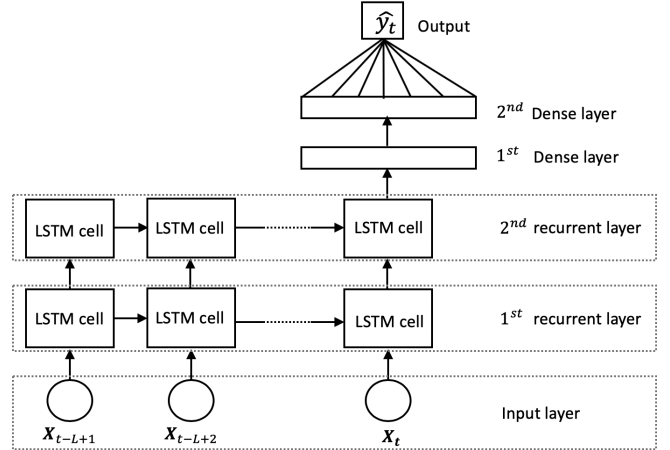


Figure 1. Unfolded representation of a general example with two LSTM layers followed by one dense layer. The input is a data sample with a sequence of data points and output is the final estimation of associated health indicator

time which also represents the lifespan of the  $i$ th device. Here, each data sample  $\mathbf{X}_t^i$  has the shape of  $L \times R$ , represented as follows:

$$\mathbf{X}_t^i = \begin{bmatrix} \mathbf{x}_{t-L+1} \\ \vdots \\ \mathbf{x}_{t-1} \\ \mathbf{x}_t \end{bmatrix} = \begin{bmatrix} x_{t-L+1,1} & x_{t-L+1,2} & \dots & x_{t-L+1,R} \\ \vdots & \vdots & \ddots & \vdots \\ x_{t-1,1} & x_{t-1,2} & \dots & x_{t-1,R} \\ x_{t,1} & x_{t,2} & \dots & x_{t,R} \end{bmatrix}$$

where  $L$  is the lag of historical data points considered in learning the current health indicator and  $R$  is the number of extracted features.  $y_t^i$  is the health indicator label at time  $t$  which is defined as degradation percentage using time information:  $y_t^i = \frac{t}{T^i}$ . For each device, the health indicator at the initial point is 0, which means there is no degradation and possibly no chance of failure, and at the end of its run-to-failure experiment is 1, which represents a high probability of failure. After data preparation, a HI estimation model will take feature data of  $\mathbf{X}$  as the input data and  $\mathbf{Y}$  as labels to learn the estimation of health indicator from the extracted features.

The HI estimation model is a composition of multiple layers of LSTM networks followed by fully connected multiple dense layers. This complex deep recurrent learning network allows the model to learn the temporal dependencies vertically and complex relationship between different features horizontally for different fault modes and degradation modes. To show how the learning model works, we take a general example of two LSTM layers followed by two dense layers as shown in Figure 1.

Because of the complexity of the learning model, we include dropout (Srivastava et al., 2014) during the training to prevent overfitting. We apply dropout to the input connections

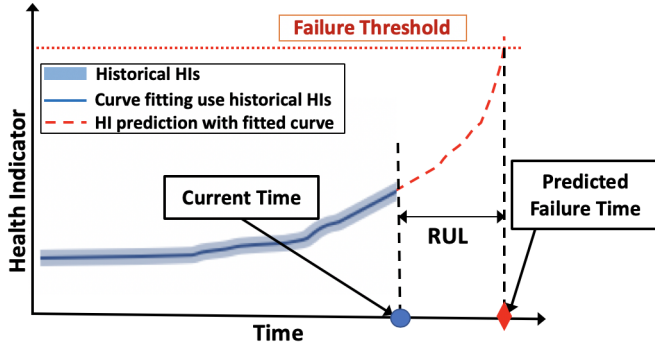


Figure 2. Simplified representation of RUL prediction with curve fitting on estimated historical health indicators

with the LSTM nodes for each LSTM layer. The dropout on the input means that the data on the input connection to each LSTM block will be excluded, at a given probability, from node activation and weight updates. In training the proposed model, mean square error (MSE) shown in Eq. 3 is used as the cost function:

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_t - \hat{y}_t)^2 \quad (3)$$

### 2.3. Remaining Useful Life (RUL) Prediction

Any equipment naturally degrades over time. RUL prediction is usually needed after the device has been operated for a certain period of time and there is a chance of failure occurrence. In the proposed method, we use the vibration measurement data from history to predict the time of end-of-life in the future. Then, the remaining useful life for the given device can be deduced.

In RUL prediction, historical health indicators are first calculated from historical vibration snapshots through the aforementioned feature extraction and trained HI estimation model. Then, a polynomial curve fitting is applied to the historical health indicators to fit the degradation mode. Finally, future health indicators can be predicted via extrapolating the fitted polynomial curve and failure is predicted to happen when the predicted health indicator reaches the failure threshold. In proposed method, the failure threshold of health indicator is set to be 1. A simplified representation of the RUL prediction is shown in Figure 2.

## 3. CASE STUDY ON IEEE PHM 2012 PROGNOSTIC CHALLENGE DATASET (PRONOSTIA DATA)

The proposed method is tested on the PRONOSTIA bearing dataset provided by FEMTO-ST Institute for remaining useful life prediction (Nectoux et al., 2012). The process is comprised of two stages: Training and Testing. In training, the input data is the sequences of vibration snapshots measured

from multiple run-to-failure bearings in learning set, and the output is the representative spectra templates for each operating condition and the trained HI estimation model.

In testing stage, the input data will be the sequence of historical vibration snapshots measured from new bearings in test set. The HI estimation model learned in training stage will be applied to estimate health indicators for the historical vibration data. Curve fitting algorithm is then applied to fit the degradation mode from historical health indicators and extrapolated to predict future health indicators. Failure is predicted to happen when the future health indicator reaches the failure threshold. Finally, the remaining useful life of the new bearings in test set can be predicted.

### 3.1. PRONOSTIA data

The PRONOSTIA data, a benchmark vibration dataset for bearing failure prognostics, was used in IEEE PHM 2012 prognostic challenge. It contains a set of real experimental data measured during the whole life span of bearings. The experimental Platform of PRONOSTIA has been extensively discussed in (Nectoux et al., 2012). There are multiple causes of bearing failure, including inner race, outer race, ball, improper lubrication, etc.. In PRONOSTIA experiments, cause of bearing failure could be one or more types of failures, which represents a real life situation. In the platform, there are two accelerometers measuring the vibration along horizontal and vertical direction respectively. In the data challenge, 3 operating conditions were considered:

- Condition 1: 1800 rpm and 4000 N
- Condition 2: 1650 rpm and 4200 N
- Condition 3: 1500 rpm and 5000 N

As shown in Table 2, six run-to-failure bearing datasets are provided to the participants as learning data to build the prognostic models. For the 11 bearings in test set, vibration data is truncated and provided till some point before failure. Thus, the task will be using the learned prognostic model to predict failure time and remaining useful life for the 11 bearings in test set. The learning set is quite small with only 2 run-to-failure bearings under each operating condition. Meanwhile, the spread of the lifespan of all bearings is wide, which varies roughly from 1.5 to 7.8 hours as shown in Table 3. Therefore, learning a good model to accurately predict RUL becomes difficult and challenging.

### 3.2. Scoring Function

To evaluate the effectiveness of proposed method, we calculate the Percent Error (PE) of RUL prediction for test bearings. PE (%Err) is defined in Eq. (4).

$$\%Err_i = 100 \times \frac{RUL_i^{real} - RUL_i^{predict}}{RUL_i^{real}} \quad (4)$$

Table 2. IEEE PHM 2012 Prognostic Challenge Dataset (Nectoux et al., 2012)

Datasets	Operating Conditions		
	Condition 1	Condition 2	Condition 3
Learning set	Bearing1_1	Bearing2_1	Bearing3_1
	Bearing1_2	Bearing2_2	Bearing3_2
Test set	Bearing1_3	Bearing2_3	Bearing3_3
	Bearing1_4	Bearing2_4	
	Bearing1_5	Bearing2_5	
	Bearing1_6	Bearing2_6	
	Bearing1_7	Bearing2_7	
	Bearing1_3	Bearing2_3	

Table 3. Learning Data in IEEE PHM 2012 Prognostic Challenge Dataset

Learning set	Operating Conditions	Lifespan (seconds)
Bearing1_1	Condition 1	28073
Bearing1_2	Condition 1	8700
Bearing2_1	Condition 2	9100
Bearing2_2	Condition 2	7960
Bearing3_1	Condition 3	5140
Bearing3_2	Condition 3	16360

where  $RUL_i^{real}$  and  $RUL_i^{predict}$  are the real and predicted RUL for the  $i^{\text{th}}$  test bearing respectively. To compare the proposed method with related studies on the same dataset, the mean and standard deviation (SD) of percent errors for all test bearings are computed.

In IEEE PHM 2012 Prognostic Challenge, underestimates and overestimates of the RUL were considered in different manners by the scoring function in Eq. (5), where  $A_i$  is the score for the  $i^{\text{th}}$  test bearing calculated from its PE ( $\%Err_i$ ). The score is 1 when PE is 0 (the predicted RUL is exactly equal to the real RUL). Non-zero PE will add penalty to decrease the score. Early prediction of failure ( $\%Err_i > 0$ , where failure is predicted to happen earlier than the actual occurrence) receives less penalty than late prediction.

$$A_i = \begin{cases} \exp[-\ln(0.5) \cdot (\%Err_i/5)] & \text{if } \%Err_i \leq 0 \\ \exp[+\ln(0.5) \cdot (\%Err_i/20)] & \text{if } \%Err_i > 0 \end{cases} \quad (5)$$

The overall score of RUL prediction result is defined as the average of scores from all test bearings, as shown in Eq. (6)

$$\text{Score} = \frac{1}{11} \sum_{i=1}^{11} (A_i) \quad (6)$$

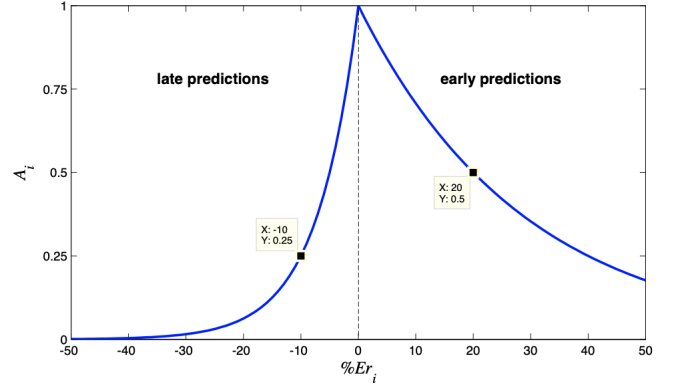


Figure 3. Scoring function of a RUL estimates according to the percent error used in IEEE PHM 2012 data challenge

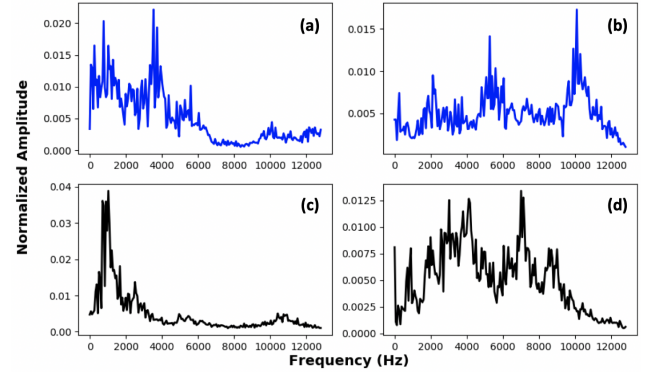


Figure 4. Normalized representative spectra for operating condition 1. There are two vibration sensors. (a) and (b) are the two spectra learned from vibration data measured when the bearing is in good condition. (c) and (d) are the two spectra associated with failure condition.

Figure 3 is the Fig. 15 in (Nectoux et al., 2012) which depicts the scoring function.

### 3.3. Representative Spectra

In PRONOSTIA bearing data, there are three operating conditions. We assume that bearings in the same operating condition have similar failure modes while bearings in different operating condition might have different failure modes. We learn the representative spectra for each of the operating conditions separately. The frequency spectrum is distributed in the frequency range up to 12.5 kHz which is half of the sampling frequency. We normalize the representative spectra before calculating the similarities between a new spectrum and the representative spectrum template. The normalized representative spectra for operation conditions 1, 2 and 3 are shown in Figures 4, 5 and 6 respectively.

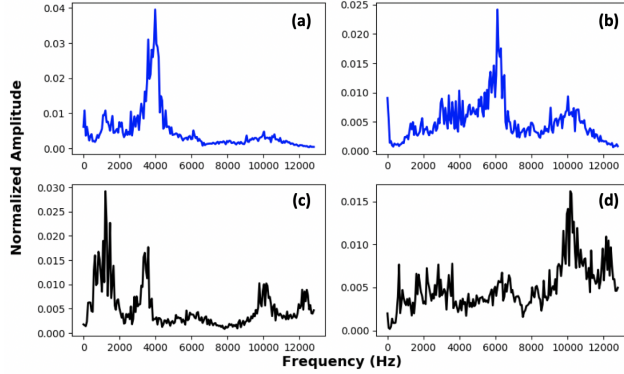


Figure 5. Normalized representative spectra for operating condition 2. Similar to Figure 4, (a) and (b) are spectra corresponding to good condition and (c) and (d) are associated with failure condition.

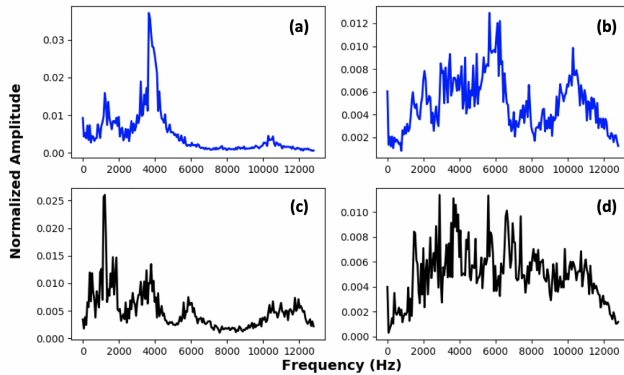


Figure 6. Normalized representative spectra for operating condition 3. Similar to Figure 4, (a) and (b) are spectra corresponding to good condition and (c) and (d) are associated with failure condition.

### 3.4. Training of HI Estimation Model

In the model training stage, hyperparameter optimization is first needed to determine a good set of hyperparameters to control the learning process for an accurate mapping from input data to the output. Grid search is a traditional hyperparameter optimization technique which scans exhaustively through a predefined grid in the hyperparameter space and finds the set of hyperparameters which produces best performance (Hsu et al., 2003). Usually the performance is measured by evaluating the model accuracy on a held-out validation set.

When implementing the proposed deep learning model for failure indicator estimation, we use grid search to determine the optimal hyperparameters which include: number of nodes in each of the two LSTM layers, number of nodes in the 1<sup>st</sup> dense layer, dropout rate, optimizer, activation function and batch size. In grid search, for each set of parameters, the model is trained six times by leaving each of the six training bearings out as a validation set. The average of the mean

square error (MSE) from validation set is used as the performance metric and we choose the set of parameters with the smallest average MSE.

The determined parameters are shown in Table 4. For the Adam optimizer, the default setting from Keras is used which has learning\_rate equals to 0.001, beta\_1 equals to 0.9 and beta\_2 equals to 0.999. Number of nodes in the 2<sup>nd</sup> dense layer is set to be 16.

Table 4. Hyperparameters for the HI Estimation Model

Hyperparameters	Value
Number of nodes in each LSTM layer	128
Number of nodes in the 1 <sup>st</sup> dense layer	128
Dropout rate	0.1
Activation function	Sigmoid
Optimizer	Adam
Batch size	128

After determining the hyperparameters for a model structure, model can be trained with learning data. In model training, learning data needs to be well split into a training set and a validation set for evaluating model's generalizability. However, the learning data in PRONOSTIA bearing data is small as mentioned in Section 3.1, which has six bearing datasets in total and each operating condition only has two bearings respectively. We use all six bearings to train an ensemble model which includes six individual models, each is trained by leaving one of the six bearings out as a validation set and the remaining five bearings as a training set. The health indicator estimated by the ensemble model is the average of health indicator values from six individual models. Ideally, health indicator is directly related to the degradation of the device. Therefore, it should have a monotonic trend for an operating device without any maintenance or repair. In the proposed method, the estimated health indicator from the ensemble model is corrected to ensure the monotonicity by interpolating each succeeding health indicator using preceding health indicator if the succeeding one is smaller than preceding one. Examples of the health indicator estimation result for bearings in learning data are shown in Figures 7, 8 and 9.

### 3.5. Testing: RUL Prediction Results for Test Set

The bearings in test set have truncated vibration data which stopped at some point before the failure happens. The task is to estimate the remaining time to the occurrence of failure based on historical vibration measurements. We apply the learned HI estimation model to bearings in test set to estimate health indicators from provided historical vibration data. By

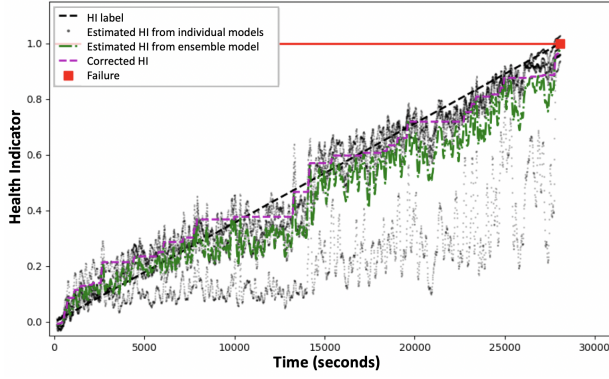


Figure 7. Failure indicator estimation result for Bearing1\_1 in learning set.

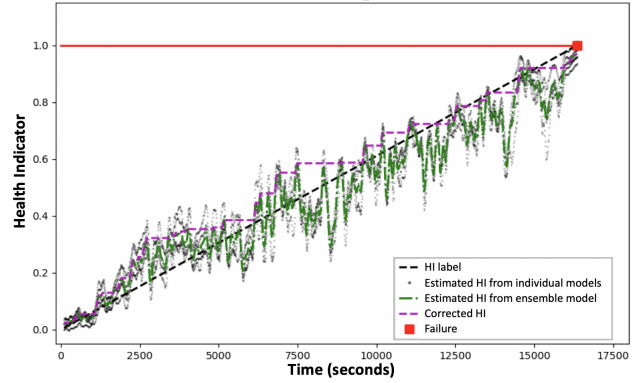


Figure 9. Failure indicator estimation result for Bearing3\_2 in learning set.

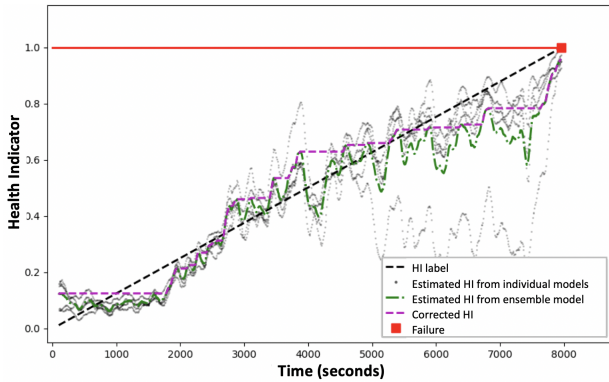


Figure 8. Failure indicator estimation result for Bearing2\_2 in learning set.

fitting the estimated historical health indicators, a predictive curve is learned which captures the historical health condition and degradation mode of the bearing. Future health indicators can be predicted using the learned predictive curve and failure is predicted to occur when future health indicator reaches a predefined failure threshold, which is set to be 1 in the case study. Finally, RUL is predicted to be the time duration between the latest point of the historical vibration data and the predicted failure point.

The estimation of historical health indicators, learned predictive curve and RUL prediction results for the bearings in test set are shown in Figure 10 for 5 bearings in operating condition 1, and Figure 11 for 6 remaining bearings in operating conditions 2 and 3. The real and predicted RUL from the proposed method for the 11 test bearings are shown in Table 5.

We also compared the proposed method with the state-of-the-art algorithms in the literature (Sutrisno et al., 2012; Lei et al., 2016; Guo et al., 2017; Chen et al., 2019) on the same data set by calculating PE of the RUL estimation for each bearing in test set as shown in Table 6. The mean and SD of PE, and the score produced using the scoring function in Eq. 5 are shown in the bottom three rows in Table 6. The proposed method

Table 5. RUL Prediction Results for Test Data

Test set	Actual RUL (seconds)	Our RUL Prediction (seconds)
Bearing1_3	5730	1496
Bearing1_4	339	118
Bearing1_5	1610	1413
Bearing1_6	1460	1412
Bearing1_7	7570	6757
Bearing2_3	7530	1399
Bearing2_4	1390	1331
Bearing2_5	3090	2967
Bearing2_6	1290	1050
Bearing2_7	580	532
Bearing3_3	820	844

produces highest score, and lowest mean PE with smallest variance.

The goal in RUL estimation is to predict the remaining time to the occurrence of failure so that maintenance or repair actions can be scheduled in time to prevent failure happens, which helps to reduce equipment downtime and improve productivity. Therefore, early prediction (failure is predicted to happen earlier than its real occurrence) is usually preferred than late prediction. As shown in Table 6, our predicted RUL is smaller than real RUL for all bearings except Bearing3\_3, indicating that the proposed method tends to make early prediction.

Similarity-based features are extracted from frequency spectrum. Time domain information from vibration measurements could help in estimating health indicator and predicting remaining useful life. Our proposed method used the combination of similarity-based features and eight statistical time domain features, as shown in Section 2.1. To demonstrate the effectiveness of the similarity-based features, we apply the process of health indicator learning and remaining useful

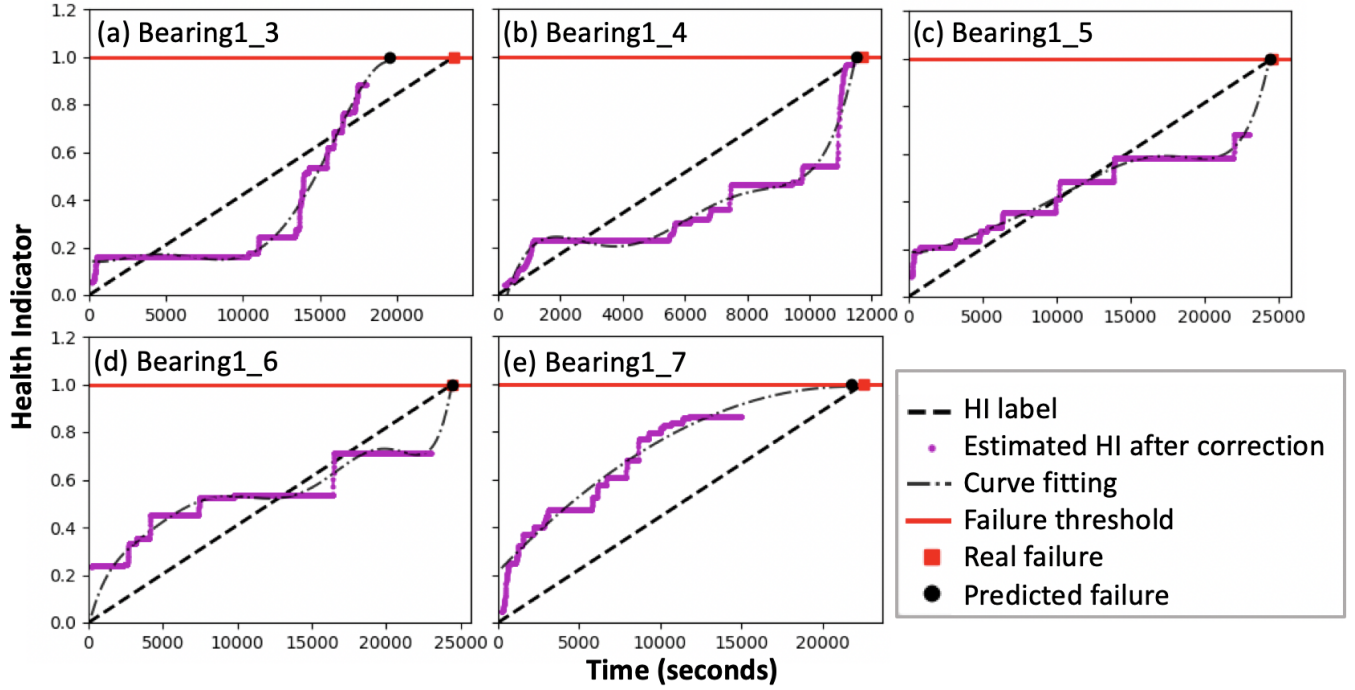


Figure 10. RUL prediction result for bearings in operating condition 1 in test set.

Table 6. Comparison of RUL Prediction Results

Test set	Percent Error (%Err)				
	Challenge Winner (Sutrisno et al., 2012)	Model-based (Lei et al., 2016)	RNN (Guo et al., 2017)	RNN with encoder-decoder (Chen et al., 2019)	Proposed method
Bearing1_3	37	-0.35	43.28	7.62	73.89
Bearing1_4	80	5.60	67.55	-157.71	65.19
Bearing1_5	9	100.00	-22.98	-72.57	12.24
Bearing1_6	-5	28.08	21.23	0.93	3.29
Bearing1_7	-2	-19.55	17.83	85.99	10.74
Bearing2_3	64	-20.19	37.84	81.24	81.42
Bearing2_4	10	8.63	-19.42	9.04	4.24
Bearing2_5	-440	23.30	54.37	28.19	3.98
Bearing2_6	49	58.91	-13.95	24.92	18.60
Bearing2_7	-317	5.17	-55.17	19.06	8.28
Bearing3_3	90	40.24	3.66	2.09	-2.93
Mean	100.27	28.18	32.48	44.49	25.89
SD	173.28	35.41	37.57	47.00	29.69
Score	0.31	0.43	0.26	0.44	0.56

life estimation to similarity-based features and time-domain features respectively. The percent errors are presented in Table 7. With the similarity-based features, the percent error and score are better than some of the related works shown in Table 6. In proposed method, the combination of time-domain features and similarity-based features from frequency

domain provides more comprehensive information and further improves the model performance.

#### 4. CONCLUSION

When estimating the remaining useful life of equipment, it is important to capture time-dependent degradation patterns



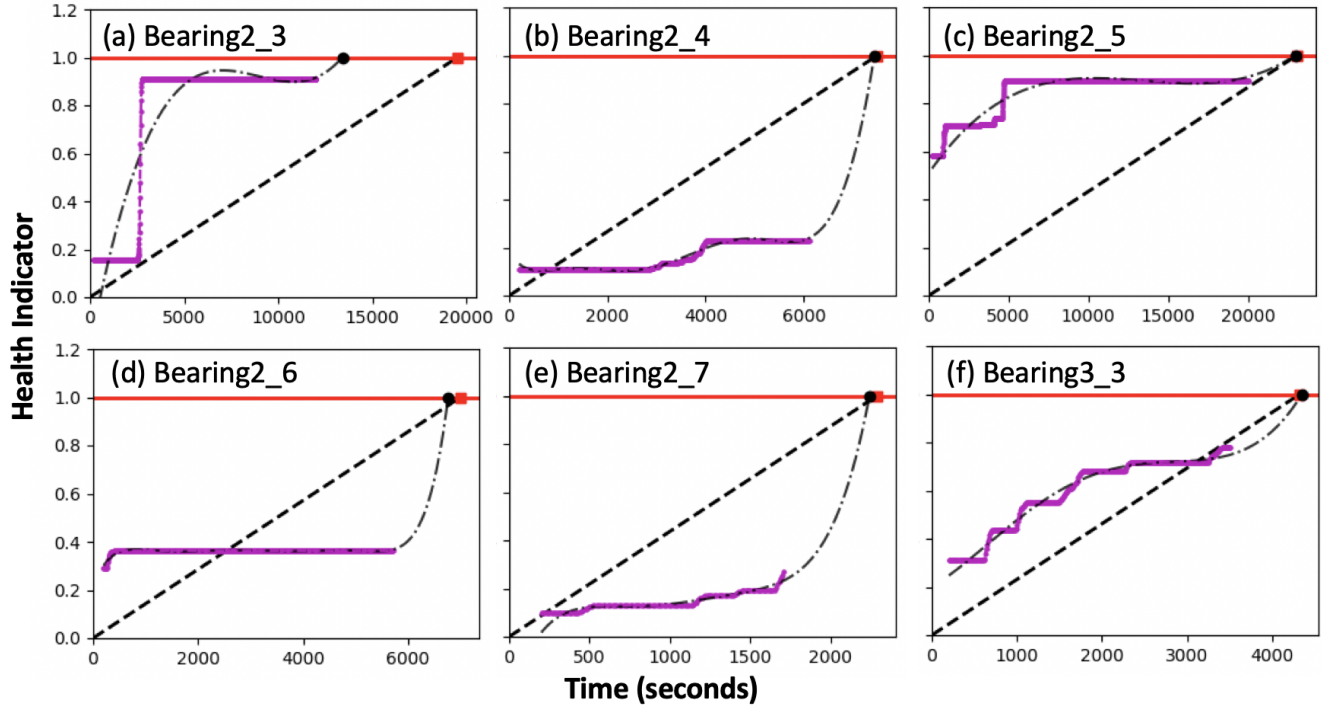


Figure 11. RUL prediction result for bearings in operating condition 2 and 3 in test set. The legend for this figure is the same as the legend in Figure 10.

Table 7. Comparison of RUL Prediction Results

Test set	Percent Error (%Err)		
	Similarity-based features	Time-domain features	Proposed method
Bearing1_3	84.00	86.83	73.89
Bearing1_4	40.12	-317.11	65.19
Bearing1_5	11.43	66.40	12.24
Bearing1_6	-5.55	21.16	3.29
Bearing1_7	81.76	94.04	10.74
Bearing2_3	31.18	83.80	81.42
Bearing2_4	75.11	-19.35	4.24
Bearing2_5	0.32	0.91	3.98
Bearing2_6	15.35	71.94	18.60
Bearing2_7	4.83	88.28	8.28
Bearing3_3	7.68	45.85	-2.93
Mean	32.48	81.33	25.89
SD	31.38	80.52	29.69
Score	0.46	0.19	0.56

from long-term sequences of measurements. In this paper, a similarity-based feature extraction method is proposed by comparing the vibration with learned representative templates in normal and failure states respectively. The similarity-based features reflect the degradation information by considering the deviation of current condition from normal and failure

condition. Then a HI estimation model is learned by training LSTM networks over the extracted features. In the case study on PRONOSTIA dataset, the result produced by the proposed method has higher accuracy compared to multiple representative works on the same dataset. The proposed method produces smallest mean percent error with smallest standard de-

viation. Furthermore, the proposed method tends to make early prediction of failure, which is good for prognostics to prevent failure occurrence.

The proposed RUL estimation method can be applied for prognostics in industries in which vibration data are measured. Vibration data are commonly measured from rotating components (such as bearings) or systems (such as assembly manipulators). The proposed method can be trained with a few failure cases and applied to predict the remaining time to failure at a given time point using historical vibration measurements.

## REFERENCES

- Al-Badour, F., Sunar, M., & Cheded, L. (2011). Vibration analysis of rotating machinery using time–frequency analysis and wavelet techniques. In (Vol. 25, pp. 2083–2101). Elsevier.
- Chen, Y., Peng, G., Zhu, Z., & Li, S. (2019). A novel deep learning method based on attention mechanism for bearing remaining useful life prediction. In (Vol. 86, p. 105919). Elsevier.
- Doguer, T., & Strackeljan, J. (2009). Vibration analysis using time domain methods for the detection of small roller bearing defects. In (pp. 23–25).
- Guo, L., Li, N., Jia, F., Lei, Y., & Lin, J. (2017). A recurrent neural network based health indicator for remaining useful life prediction of bearings. In (Vol. 240, pp. 98–109). Elsevier.
- Hsu, C., Chang, C., Lin, C., et al. (2003). A practical guide to support vector classification. Taipei.
- Lei, Y., Li, N., Gontarz, S., Lin, J., Radkowski, S., & Dybala, J. (2016). A model-based method for remaining useful life prediction of machinery. In (Vol. 65, pp. 1314–1326). IEEE.
- Nectoux, P., Gouriveau, R., Medjaher, K., Ramasso, E., Chebel Morello, B., Zerhouni, N., & Varnier, C. (2012, 06). Pronostia: An experimental platform for bearings accelerated degradation tests. In (pp. 1–8).
- Plante, T., Nejadpak, A., & Yang, C. (2015). Faults detection and failures prediction using vibration analysis. In *2015 IEEE Autotestcon* (pp. 227–231).
- Singhal, A., & Khandekar, M. (2013). Bearing fault detection in induction motor using fast fourier transform. In *IEEE Int. Conf. on advanced research in engineering & technology*.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. In (Vol. 15, pp. 1929–1958). JMLR. org.
- Sutrisno, E., Oh, H., Vasan, A. S. S., & Pecht, M. (2012). Estimation of remaining useful life of ball bearings using data driven methodologies. In *2012 IEEE conference on prognostics and health management* (pp. 1–7).
- Zhang, S., Zhang, S., Wang, B., & Habetler, T. (2020). Deep learning algorithms for bearing fault diagnostics—a comprehensive review. In (Vol. 8, pp. 29857–29881). IEEE.