KSME

Springer

Journal of Mechanical Science and Technology 38 (5) 2024

Original Article
DOI 10.1007/s12206-024-0402-8

Keywords:
· Bearing
· Degradation
· LSTM
· Modified multiscale
· Permutation entropy
· RUL

Correspondence to:
Prashant Kumar Sahu
Prashantphd1@iitkgp.ac.in

Citation:
Sahu, P. K., Rai, R. N. (2024). LSTM-based deep learning approach for remaining useful life prediction of rolling bearing using proposed C-MMPE feature. Journal of Mechanical Science and Technology 38 (5) (2024) 2197~2209.
http://doi.org/10.1007/s12206-024-0402-8

Received July 4th, 2023
Revised January 3rd, 2024
Accepted January 20th, 2024

† Recommended by Editor No-cheol Park

LSTM-based deep learning approach for remaining useful life prediction of rolling bearing using proposed C-MMPE feature

Prashant Kumar Sahu and Rajiv Nandan Rai

Subir Chowdhury School of Quality and Reliability, Indian Institute of Technology, Kharagpur, India

Abstract Prognostic health management (PHM) is essential for the predictive maintenance of industrial systems, aiming to predict the remaining useful life (RUL) of system to ensure safe, reliable, and cost-effective operation of the machinery. This work proposes an innovative method for RUL prediction of bearings, by combining a health indicator (HI) proposed from the absolute cumulative modified multiscale permutation entropy (C-MMPE) feature with a deep learning long short-term memory (LSTM) model. The work also introduces a virtual health degree for bearings, using an exponential degradation pattern as the target function for the LSTM model output. Experimental validation showcases the effectiveness of proposed approach, achieving a high score value of 0.81 and demonstrating a lower mean absolute error value of 7.38 in RUL prediction for test bearings compared to conventional features and regression labeling functions. This highlights the superior RUL prediction capability of the proposed methodology.

1. Introduction

Rolling bearing forms an essential part of rotating machineries, such as motors, pumps, conveyors, gearboxes, etc., as it determines their smooth functioning [1-3]. It constantly runs under hostile environments, varying loads and temperatures leading to catastrophic failures of the components/ systems if the faults remain undetected [4]. The prognostic health management (PHM) of rolling bearings is essential for the predictive maintenance of the mechanical systems by predicting the RUL to maintain safe and economical operation and prevent an unexpected shutdown in the industries [5]. The data-driven approach for PHM for any machinery includes three phases: feature extraction from condition monitoring (CM) data, degradation monitoring, and RUL prediction [6, 7]. Sensors are deployed to extract useful CM data that provides information about the system's health. An accelerometer, used to capture the vibration signal, is the most effective fault monitoring technique for rolling bearings, showing dynamic changes in the characteristic signal on the occurrence of faults [8]. In vibration analysis, the features are an important factor that reflects the degradation behavior of the system.

To monitor the health of bearing state, it is crucial to extract useful information from the vibration signals in terms of features. Generally, time-domain features such as root mean square (RMS) value, kurtosis, skewness, peak-to-peak value etc., are widely used as health indicators (HI) for bearings [9]. Spectral features such as power spectral density, spectral kurtosis, spectral Skewness etc., represent the system performance characteristics in the frequency domain [10]. These standard features are sensitive to a particular failure mode and fail to describe the overall degradation process of the bearing. The vibration signals excited by local defects of bearings show non-stationary and non-linear behavior due to the presence of external phenomena such as strikes, external noise components, friction, etc. Therefore, for identifying the dynamic non-linear features of bearing, HI construction has attracted the researcher's attention [11, 12]. Entropy approaches are frequently used to assess the non-stationary and non-linear dynamic characteristics of the time series data. Therefore, using entropy theory to assess the overall complexity of the bearing vibration signal enables the evaluation of the degradation state [13]. The evolution in entropy methods for information measurement is explained in the subsequent paragraph.

Shannon [14] developed the concept of information entropy and explained the problem in the information measurement but did not describe the information obtained from a change of signal. Later, approximate entropy, sample entropy, and permutation entropy concept were developed based on the information entropy theory. Pincus developed approximate entropy (ApEn) to measure the complexity in short finite time series [15]. However, ApEn undergoes a similarity problem and shows poor consistency in the entropy calculation. To improve the ApEn method, Richman proposed a method named as sample entropy (SpEn) [16], which has better characteristics for shorter data and shows better consistency when compared to ApEn, but it has low computational efficiency. Later, permutation entropy (PE), a novel technique, was developed by Bandt et al. to assess the complexity of the dynamic behaviour of non-stationary and non-linear time series [17]. Several works have employed methods like PE, fuzzy entropy, and dispersion entropy to evaluate the condition of rolling bearings [18-21]. Dispersion entropy, which relies on the spread of signal values within a time window, may not effectively capture localized, transient anomalies or variations crucial for assessing bearing health. Consequently, it might lack sensitivity to specific types of bearing faults or fail to provide early warnings for potential issues. The computation of fuzzy entropy can be resource-intensive, particularly for extensive datasets or high-dimensional data. In contrast, permutation entropy stands out for its superior performance, computational efficiency, sensitivity, and remarkable resistance to noise when compared to other entropy features [22]. To minimize further loss of information, multiscale permutation entropy (MSPE) was developed to enhance the effectiveness of PE algorithm by fusing the PE and the multiscale technique concept [23]. Many studies have been found using the MSPE algorithm for bearing fault diagnosis by assessing the complexity of the vibration signals [24], [25]. The multiscale coarse-graining (CG) procedure for entropy calculation significantly reduces the data point length and can lead to an inaccurate entropy value for short time series. The moving average graining (MAG) procedure in multiscale entropy is introduced by Wu et al. to confront the data length problem in the CG process to construct a new time series sequence [26]. The authors concluded that the MAG reflects better long-range correlations of a short-term time series. Therefore, MAG in multiscale entropy provides more accurate entropy values for short time series data. They implemented the MAG on sample entropy calculation to detect the bearing fault. In this work, modified multiscale permutation entropy (MMPE) is utilized which is formed by combining PE, the multiscale, and the MAG approach to construct a HI for the health assessment of rolling bearing. It has a negligible effect on new time series sequence length [27]. The features effectiveness is measured by investigating performance metrics such as monotonicity, robustness, and trendability [28]. It is observed that the MMPE is a dominant feature and has been selected for further analysis. The next step is to estimate an absolute cumulative effect of features. The vibration features are affected by numerous factors, such as noise, friction, strike, etc., that are visible in the form of some local fluctuations and cause non-ideal behavior to represent the machinery degradation process. A bearing deterioration is the cumulative effect of all processes. As a result, it is vital to comprehend the degrading progression from a cumulative aspect. The continuous accumulation of vibration features data from the cumulative aspect carries enriched prior information, decreases local fluctuations, and generates a more reliable trend characteristic [29]. Sahu and Rai proposed a degradation monitoring and RUL prediction technique for rolling bearings using the C-MMPE feature. They found that the C-MMPE is an effective feature that is sensitive to an incipient fault and precisely predicts RUL using an exponential degradation model [30]. Considering the advantage of C-MMPE, this feature is selected for the construction of HI for regression analysis to further reveal its effectiveness for intelligent RUL prediction using deep learning techniques.

In regression analysis, it is necessary to define the output target function, i.e., to form a virtual life or health degree, representing the degradation behaviour of the bearing. This virtual life or health degree is used as an output target function for regression analysis for RUL prediction. Bearing health degree is constantly changing in its life cycle. Thus, a precise representation of the degradation trend is important to track the bearing health and predict its RUL. Therefore, it is crucial to create a labelling function to define the health degree or virtual health for bearing [31]. The traditional methods simulate the bearing's life cycle pattern as a linear or piecewise function [32]. However, these functions may not characterize the actual life scenario of bearing. Thus, there is scope for the selection of some other target function to represent the virtual life of bearing. In this work, linear, piecewise, quadratic, and exponential functions are considered to represent virtual life. After data processing is done, the next step is to utilize a proper deep-learning tool for RUL prediction, as explained in the subsequent paragraph.

The remaining useful life of any machinery is defined as the time length of the system from a current state to a failed state [33]. Traditionally, failure models have been developed to estimate the RUL of equipment by analyzing the product degradation mechanisms [34]. However, this approach needs a lot of experience and complex modeling of failure mechanisms. Data-driven prediction techniques utilize prior or historical data and have gained huge popularity due to the development of deep learning techniques [35]. Deep learning techniques can quickly and efficiently extract useful information from huge amounts of data because of their strong information extraction capabilities. Hence, researchers prefer a deep learning-based RUL prediction model [36, 37]. Long short-term memory (LSTM) is a deep learning technique that can efficiently deal with sequential data. Lei et al. utilize the LSTM for CM and fault diagnosis of a wind turbine [38]. The LSTM model showed performance superiority compared to support vector machine (SVM), recurrent neural network (RNN), multi-layer perceptron (MLP), and convolution neural network (CNN). The LSTM model is also widely used for direct RUL predictions. Various studies have been carried out for RUL prediction of an aircraft engine using the LSTM model [39-41]. Mao et al. used the LSTM directly to predict the RUL of bearings by using the vibration signal's extracted features as input [42]. Rathore et al. proposed a model for extractive prognostic feature by developing transfer learning based bi-LSTM network [43]. This work utilizes the LSTM deep learning techniques by considering the advantages of LSTM for direct prediction of RUL from extracted features. The above studies avoided the complex feature extraction and selection procedure, utilizing traditional features that are sensitive to failure modes. Secondly, as discussed in the LSTM model above, most papers utilize the linear index corresponding to the RUL that may not characterize the actual degradation behavior of bearings.

Based on the limitations and gaps in the research literature on RUL prediction methodologies of bearings, as discussed in the preceding paragraphs, this paper aims to develop a single dominant HI and defines the effective virtual life of bearing for RUL prediction using the deep learning LSTM model. The features are extracted from the vibration signals, followed by the feature performance measurement to reflect the effectiveness of each feature. Then, each selected feature is normalized, and its absolute cumulative effect is computed to form HI. The exponential output target function is subsequently defined to represent the virtual life or health degree of bearings. Finally, the LSTM model is implemented for direct RUL prediction from the extracted features. The concept of direct RUL prediction using LSTM has mainly been taken into consideration with traditional time domain features with linear or piecewise functions. As a result, the novelty of this paper lies in considering novel HI based on C-MMPE and exponential output target function with LSTM model for RUL prediction. The obtained results from the proposed methodology indicate that the HI constructed from C-MMPE is a dominant and sensitive feature that characterizes the degradation process precisely in the rolling bearing and more accurately predicts the RUL with exponential target function compared to other features and target functions.

This paper is organized as follows: Sec. 2 briefly describes the detailed procedure of the proposed methodology. In Sec. 3, a dataset description is provided. Sec. 4 evaluates the proposed method on experimental datasets. Finally, Sec. 5 concludes the work.

2. Proposed methodology

The proposed method flowchart is illustrated with the help of Fig. 1. The bearings vibration signals are captured with the help of an accelerometer. The vibration analysis is performed to predict the RUL of the bearing. Firstly, the virtual life or health degree is defined for the bearing. Then exponential health degree is proposed and set as the output target function for RUL prediction. Further, the vibration data is processed, including feature extraction, selection, and construction of HI, and subsequently considers the degradation of bearing. The performance of each feature is measured by calculating the monotonicity, trendability, and robustness. After HI construct, the LSTM regression model is trained with training datasets to predict the RUL of test bearing.

The steps involved in HI construction are explained in detail in the subsequent subsections:

2.1 Health indicator construction

This work considers five traditional vibration features, such as RMS, kurtosis, skewness, spectral skewness and spectral kurtosis and five entropy-based features, such as PE, MSPE, MMPE, dispersion and fuzzy entropy to develop the HI that reflects the degradation behavior of bearings. Notably, PE and its advanced variations have proven to be particularly effective in serving as health indicators for bearings compared to dispersion and fuzzy entropy. The algorithm for dispersion and fuzzy entropy are briefly outlined in Refs. [44, 45]. The development and significance of permutation entropy-based fea-

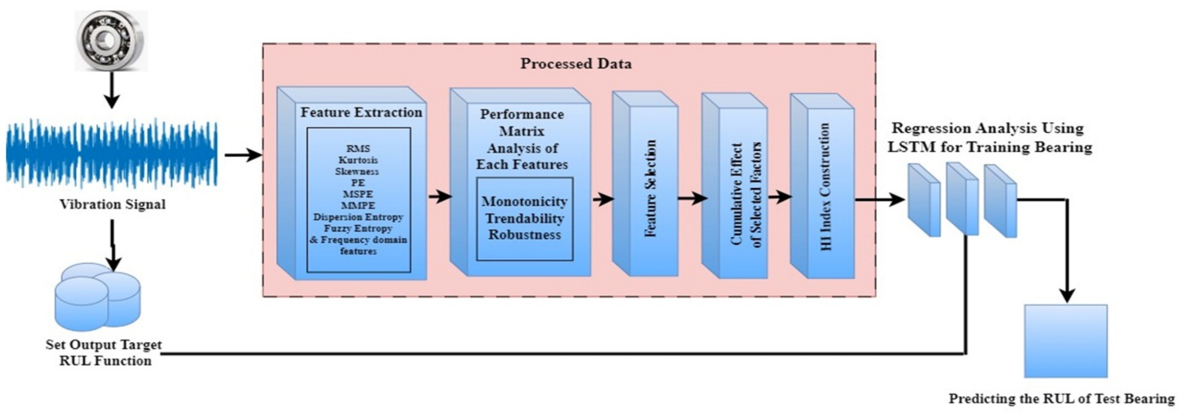

Fig. 1. Proposed methodology for RUL prediction of bearing.

tures are elaborated in the subsequent subsections.

2.1.1 Permutation entropy

Bandt and Pompe introduced the permutation entropy concept in 2002 [17]. It is widely used in several different domains to assess the complexity of time series data.

For a given signal, $x = x_1, x_2, x_3, \ldots\ldots, x_N$ of length 'N', the first step is to form a matrix of overlapping column vectors by splitting a one-dimensional time series signal data using hyperparameters 'm' & 'τ'.

$$X_i^m = \left\{x_i, x_{i+\tau}, \ldots, x_{i+(m-2)\tau}, x_{i+(m-1)\tau}\right\},$$
$$i = 1, 2, \ldots\ldots, i+(m-1)\tau .$$

Where 'm' and 'τ' represents the embedding dimension and time lag, respectively, these hyperparameters determine the amount of information each vector holds.

Next, embedding vector 'X_i^m' is rearranged in increasing order.

$$\left\{x_{i+(r_1-1)\tau}, x_{i+(r_2-1)\tau}, \ldots, x_{i+(r_m-1)\tau}\right\}.$$

There will be '$m!$' different possible ordinal permutations in 'm' dimensional space. The relative frequency for each permutation 'π' is determined using the following formula:

$$p(\pi) = \frac{Number\ \{i \mid i \le N-(m-1)\tau,\ x_i^m\ has\ type\ \pi\}}{N-(m-1)\tau} . \quad (1)$$

Finally, PE is defined as follows:

$$H_{PE}(m) = -\sum p(\pi)\ln\left(p(\pi)\right) \quad (2)$$

when $H_{PE}(m) = \log(m!)$, the $H_{PE}(m)$ reaches its maximum value indicating that the permutation probabilities of all possible ordinal patterns are the same. Finally, the normalization of PE is performed for easier comparison and interpretation.

$$H_{NPE}(m) = \frac{H_{PE}(m)}{\ln(m!)} \quad (3)$$

where, $0 \le H_{NPE}(m) \le 1$.

When $H_{NPE}(m) = 0$ denotes extremely periodic signals, whereas $H_{NPE}(m) = 1$ denotes that all ordinal patterns have the same probability.

2.1.2 Multiscale permutation entropy (MSPE)

Costa et al. proposed a multiscale analysis to extract more dynamic information than a single scale [23]. This method introduced a scale factor to divide and generate the new sequence, but it significantly impacts the time series length. The

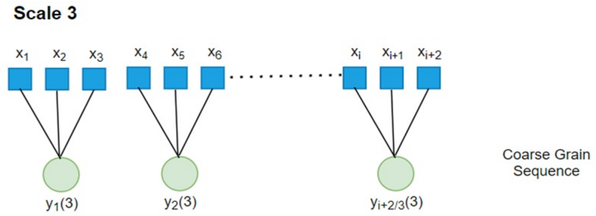

Fig. 2. Illustrate the CG procedure to generate a new sequence of time series with scale factor s = 3.

MSPE approach uses the following two steps:

1) First, the non-overlapping window of time series data are formed by splitting an original time series $\{x_i, i = 1, 2, 3, \ldots, N\}$ with a scale factor of length 's', to generate a new CG time series 'y_j^s'. Fig. 2 illustrates the CG method. The equation to generate a CG time series sequence is given as follows:

$$y_j^s = \frac{1}{s}\sum_{i=(J-1)s+1}^{js} x_i, \quad 1 \le j \le \frac{N}{s} . \quad (4)$$

2) Finally, the MSPE is obtained by calculating PE for new sequence for specified scale factor 's' and embedded dimension 'm'.

$$MSPE(x,s,m,\tau) = PE\left(y_j^s, m, \tau\right) . \quad (5)$$

2.1.3 Modified multiscale permutation entropy (MMPE)

Modified multiscale permutation entropy (MMPE) is utilized in this work, which is formed by combining PE, the multiscale, and the MAG approach to construct a HI for the health assessment of rolling bearing. The moving-average graining (MAG) method is more dependable and noise-sensitive than the CG method for short-term time series analysis as it doesn't affect the length of the new time series sequence, making it more sensitive to incipient fluctuation. The proposed MMPE approach uses the following two steps:

1) To reflect the dynamic behaviour of the signal, a new time series is formed 'y_j^s' from the original time series data $\{x_i, i = 1, 2, 3, \ldots, N\}$ by applying the MAG with a scale factor of length 's'. Fig. 3 illustrates the MAG method. The equation to generate a moving average time series sequence is given as follows:

$$y_j^s = \frac{1}{s}\sum_{i=j}^{j+s-1} x_i, \quad 1 \le j \le N-s+1 . \quad (6)$$

2) Now, the MMPE is determined by measuring the PE for the new sequence with the specified embedded dimension 'm' and scale factor 's'.

$$MMPE(x,s,m,\tau) = PE\left(y_j^s, m, \tau\right) . \quad (7)$$

Once all features are obtained, the next step is to construct

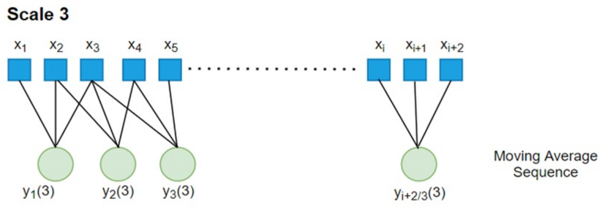

Fig. 3. Illustrate the MAG procedure to generate a new sequence of time series with scale factor s = 3.

HI using the following equation given below:

$$HI = 1 - Normalized \left(\text{absolute cumulative effect of } features\right). \tag{8}$$

The continuous accumulation of vibration features data from the cumulative aspect carries enriched prior information, decreases local fluctuations, and generates a more reliable trend characteristic. The above equation is used to construct a HI such as it forms a decreasing trend and shows better monotonic trend to the output target function.

2.2 Feature evaluation

Feature reflects the health of bearing that is used for fault diagnosis and prognosis. Some features are sensitive to particular failure modes and unsuitable for RUL prediction. Three performance indicators are used, monotonicity (Mon), trendability (Tre), and robustness (Rob), to screen the features that can efficiently represent the degradation process and are further used for the predictability.

The absolute difference between each feature's number of positive and negative derivatives determines the monotonicity, and its range varies between 0 and 1. The higher monotonicity value represents the better fitness of the feature. The trendability scale runs from 0 to 1, and the greater the trend index, the more linearly the feature sequence is correlated with time (t). The robustness scale also varies between 0 and 1. Robustness represents the fluctuation in the features. The smaller robustness value indicates that, the more the feature fluctuates, resulting in greater uncertainty.

The mathematical expression for monotonicity, trendability, and robustness [28, 46] are placed below as Eqs. (9)-(11):

$$Mon(f) = \left| \frac{\#\, of\, \frac{d}{df} > 0}{k-1} - \frac{\#\, of\, \frac{d}{df} < 0}{k-1} \right| \tag{9}$$

$$Tre(f,t) = \frac{\left| k\sum_i f_i t_i - \sum_i f_i \sum_i t_i \right|}{\sqrt{\left[k\sum_i f_i^2 - (\sum_i f_i)^2\right]\left[k\sum_i t_i^2 - (\sum_i t_i)^2\right]}} \tag{10}$$

$$Rob(f) = \frac{1}{k}\sum_i \exp\left(-\frac{f_i - \overline{f_i}}{f_i}\right). \tag{11}$$

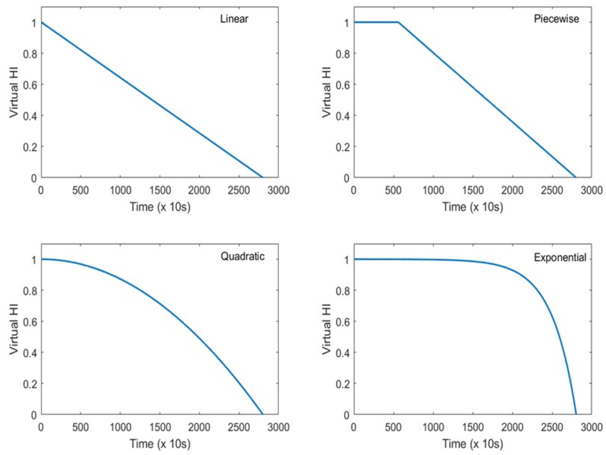

Fig. 4. Different functions curve for bearing virtual RUL.

Where f represents the original feature, k is the number of observations in a particular feature, d/df is the average difference of the fraction of derivatives for each feature, t is the time index and $\overline{f}$ represents the smoothing processing of the original feature.

In some instances, a specific metric may exhibit a slight advantage over others, influencing the feature selection process. To ensure the consistent and accurate selection of the most suitable degradation features suitability can be measured using a single evaluation metric. A linear weighted comprehensive indicator (CI) is proposed as a single evaluation metric to evaluate the feature more thoroughly. Monotonicity is assigned the highest weightage value due to its paramount significance in the feature selection process. Its prominence lies in the ability to create simpler and more interpretable models. In research, where model transparency holds utmost importance, giving priority to features with a well-defined monotonic relationship is of the highest priority.

The CI is defined as follows [31]

$$CI = 0.4{*}Mon(f) + 0.3{*}Tre(f,t) + 0.3{*}Rob(f). \tag{12}$$

Features with a high CI value indicate better degradation behavior of bearing.

2.3 Virtual RUL construction

The health state of the bearing constantly changes during its life cycle. A suitable representation of the virtual degradation trend or health degree is essential to represent the health state of the bearings for precise RUL prediction.

Therefore, it is essential to design an output labelling function for regression analysis describing the bearings health degree or virtual life. In this work, an exponential function is proposed as a health degree and compared with linear, piecewise, and quadratic functions to show its superiority. The formulas for the construction of these labelling functions are mentioned in Table 1. In Fig. 4, the graphical representations of these function

Table 1. Function curves formula to represent the virtual RUL for bearing.

Sl. No	Function curve	Formula	
1	Linear	$f(t) = -\left(\frac{1}{t_n} * t_i\right) + 1$	(13)
2	Piecewise	$f(t) = \begin{cases} 1, t_i \le t_j \\ \left(\frac{1}{t_j - t_n} * t_i\right) + \left(\frac{t_n}{t_n - t_j}\right), t_i > t_j \end{cases}$	(14)
3	Quadratic	$f(t) = -\left(\frac{1}{t_n^2} * t_i^2\right) + 1$	(15)
4	Exponential	$f(t) = d - \exp(t\tau + a)$ $\begin{cases} f(t_{min}) = 1 \\ f(t_{max}) = 0 \end{cases}$ Where a = convergence rate hyperparameter d and τ can be determined by solving the above two equations.	(16)
Where, t_n is the whole life duration of bearing, t_i is the current time and t_j is the initial degradation time of bearing.			

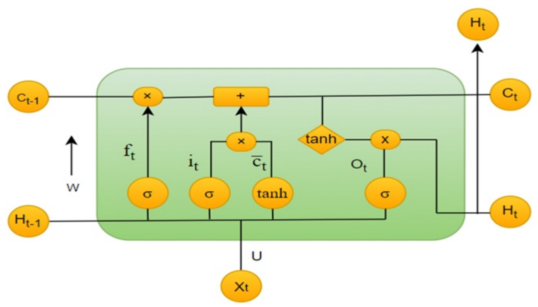

Fig. 5. LSTM network architecture.

curves are shown.

2.4 LSTM network

The architecture of the LSTM network is shown in Fig. 5. The reason behind the selection of the LSTM network lies in its ability to solve vanishing gradient problems and long-term dependency in RNN [47]. The architecture of LSTM consists of three gates named input, output, and forget gates, as shown in Fig. 5. It has both short- and long-term memory. The information is passed through the network and retrieved at a very later state to identify the context of prediction [48].

Mathematically, the LSTM network can be expressed as follows:

Forget gate:

$$f_t = \sigma\left(X_t U^f + H_{t-1} W^f\right) . \quad (17)$$

Input gate:

$$i_t = \sigma\left(X_t U^i + H_{t-1} W^i\right) . \quad (18)$$

Table 2. Pseudocode for RUL prediction.

Pseudocode for RUL prediction of Bearing Using Constructed HI and LSTM Network
1: import library
2: Load dataset and extract features
3: Construct HI   for feature in features: construct_health_index (feature)
4: Split HI into train and test sets
5: Create LSTM model   model = create_lstm_model (input_units=1, lstm_units=50, output_units=1, optimizer='adam')
6: Train LSTM model   for epoch in range (epochs = 400):   train_lstm_model (model, train_data, learning_rate=0.001, batch_size=256)
7: Make Prediction   predictions = model.predict (test_data)
8: Calculate Score value and ER%

Output gate:

$$O_t = \sigma\left(X_t U^o + H_{t-1} W^o\right) . \quad (19)$$

Cell state:

$$\bar{C}_t = \tanh\left(X_t U^g + H_{t-1} W^g\right) . \quad (20)$$

Updated cell state:

$$C_t = \sigma\left(f_t * C_{t-1} + i_t * \bar{C}_t\right) . \quad (21)$$

Output:

$$H_t = \tanh(C_t) * O_t . \quad (22)$$

Where the previous LSTM cell output is represented by H_{t-1} and its cell state by C_{t-1}. The LSTM unit input vector is denoted by X_t. U and W represent the input and the recurrent weight matrix for the gate denoted by $t* \in \{i, f, g, o\}$. In the process of network training, these parameters are learned and updated. The sigmoid and tangent hyperbolic activation functions are represented by σ and $\tanh$, respectively. Based on the previous state C_{t-1} and the input gate i_t, the LSTM cell can update the weights according. The gating mechanism, which is the primary characteristic of the LSTM cell, is responsible for measuring the capability of the input signals over long-interval dependency [49]. The proposed method maintains the following LSTM parameters consistent across all Health Indicators (HI) to optimize the virtual remaining useful life (RUL) and enhance bearing RUL prediction:

- Learning rate: 0.001
- Batch size: 256
- Number of epochs: 400

These parameters have been carefully chosen and kept constant to ensure the most effective virtual RUL and HI construction for improved bearing RUL prediction.

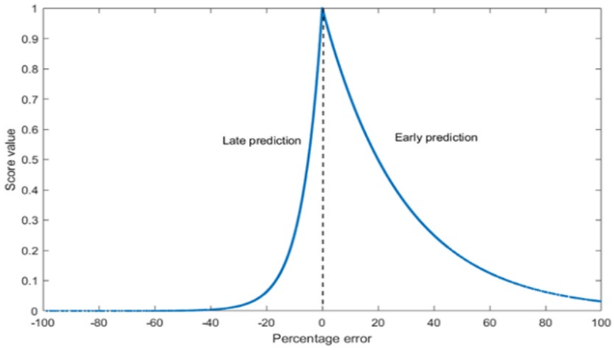

Fig. 6. Score function curve for RUL prediction.

The pseudocode is presented in Table 2 illustrate the use of methodology for RUL prediction of bearing.

2.5 Prediction performance evaluation

The error percentage (ER%) is calculated to validate the effectiveness of the proposed method for RUL prediction of test bearings. ER% is defined in Eq. (23).

The score function was introduced in PHM 2012 prognostic challenge to underestimate and overestimate the RUL prediction, as stated in Eqs. (24) and (25). A_i is the score for the i^{th} test bearing calculated from its $ER\,\%$. When the $ER\,\%$ is 0, the score value is 1, signifying that the predicted RUL is equal to the actual RUL. If the $ER\,\%$ is non-zero, then a penalty is added to decrease the score. When $ER_i\% > 0$ indicates the early failure prediction of the system and receives less penalty compared to late prediction.

$$ER\% = \frac{RUL_{Actual} - RUL_{Predicted}}{RUL_{Actual}} * 100 \tag{23}$$

$$A_i = \begin{cases} \exp\left[-\ln(0.5) * \left(\frac{ER_i\%}{5}\right)\right] & if\ ER_i\% \le 0 \\ \exp\left[+\ln(0.5) * \left(\frac{ER_i\%}{20}\right)\right] & if\ ER_i\% > 0\,. \end{cases} \tag{24}$$

The overall RUL prediction score is determined by averaging the score value results of all test bearings given by:

The score value is utilized as an evaluation index to evaluate the underestimation and overestimation of the predicted RUL, as shown in Fig. 6. Further, to evaluate the accuracy of the proposed method, the mean and absolute average of ER% are utilized.

$$Score = \frac{1}{11}\sum_{i=1}^{11} A_i \tag{25}$$

$$\overline{ER\%} = \frac{1}{N}\sum_{i=1}^{N} ER_i\,\% \tag{26}$$

$$|\overline{ER\%}| = \left|\frac{1}{N}\sum_{i=1}^{N} |ER_i|\,\%\right|\,. \tag{27}$$

Table 3. Experimental operating conditions for PRONOSTIA bearing test rig.

Datasets	Operating conditions		
	1800 rpm, 4000 N	1650 rpm, 4200 N	1500 rpm, 5000 N
Training set	Bearing 1_1	Bearing 2_1	Bearing 3_1
	Bearing 1_2	Bearing 2_2	Bearing 3_2
Test set	Bearing 1_3	Bearing 2_3	Bearing 3_3
	Bearing 1_4	Bearing 2_4	
	Bearing 1_5	Bearing 2_5	
	Bearing 1_6	Bearing 2_6	
	Bearing 1_7	Bearing 2_7	

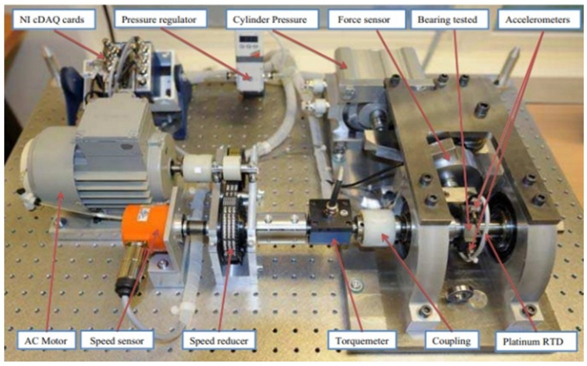

Fig. 7. The experimental PRONOSTIA platform for accelerated bearing degradation tests [50].

3. Dataset description

This paper performs the experimental validation on the PRONOSTIA platform-bearing dataset provided by the FEMTO-ST Institute in PHM 2012 [50]. The illustration of this bearing test platform is presented in Fig. 7. The accelerometer sensors are mounted on the bearing outer ring to capture the vibration signals in both horizontal and vertical directions.

The raw vibration signals are captured at every 10 seconds interval, and each recording lasts for 0.1 seconds with a sampling frequency of 25.6 kHz. The experiment is conducted at a constant rotational speed and payload conditions such as 1800, 1650, and 1500 rpm at 4000 N, 4200 N, and 5000 N, respectively. The bearing is considered to work normally in the experiment if the vibration signal amplitude is less than 20 g. The dataset consists of 6 training and 11 test datasets under three different working conditions, as shown in Table 3.

4. Results and discussion

The selection of the MMPE feature for RUL prediction is justified by calculating the comprehensive indicator (CI) value as explained in Eq. (12). The CI is calculated for test and training bearings for all six selected features, such as MMPE, PE, MSPE, RMS, skewness, and kurtosis as shown in Tables 4 and 5, respectively. The CI evaluation value for MMPE and

Table 4. Comprehensive Indicator for test bearings.

Features	Test bearings											
	1_3	1_4	1_5	1_6	1_7	2_3	2_4	2_5	2_6	2_7	3_3	Mean
MMPE	0.5358	0.5036	0.4996	0.4802	0.5421	0.4339	0.3841	0.3976	0.4473	0.4421	0.5745	0.4764
PE	0.3909	0.3167	0.5337	0.5597	0.5283	0.4755	0.5986	0.467	0.5893	0.4803	0.2984	0.4762
MSPE	0.3271	0.4035	0.3639	0.366	0.3176	0.3114	0.3946	0.3395	0.3015	0.3924	0.3756	0.3539
Dispersion entropy	0.4077	0.4107	0.351	0.3186	0.3003	0.3114	0.3394	0.3369	0.4469	0.3114	0.3345	0.3517
Fuzzy entropy	0.4255	0.3879	0.3239	0.339	0.2912	0.3108	0.3564	0.3192	0.3679	0.3517	0.3786	0.3502
Rms	0.3492	0.3178	0.3228	0.2988	0.3809	0.2232	0.3061	0.4421	0.2656	0.2864	0.4281	0.3292
Spectral skewness	0.3076	0.3169	0.3148	0.3612	0.3247	0.3066	0.2706	0.3582	0.2828	0.3133	0.4382	0.3268
Spectral kurtosis	0.4005	0.3313	0.4489	0.3136	0.3025	0.2914	0.2537	0.2521	0.2491	0.398	0.3258	0.3243
Kurtosis	0.3257	0.374	0.3518	0.2766	0.3941	0.1628	0.2679	0.2833	0.3963	0.229	0.2946	0.3051
Skewness	0.2415	0.1326	0.1766	0.1453	0.2875	0.1176	0.1535	0.1367	0.319	0.2049	0.2396	0.1959

Table 5. Comprehensive Indicator for training bearing.

Features	Training bearings						
	1_1	1_2	2_1	2_2	3_1	3_2	Mean
MMPE	0.4582	0.5873	0.4951	0.5319	0.4115	0.4176	0.4836
PE	0.4584	0.4925	0.3549	0.3219	0.5757	0.5123	0.4526
MSPE	0.3701	0.3312	0.3761	0.4723	0.3197	0.3745	0.3740
Dispersion entropy	0.4702	0.4609	0.2786	0.3760	0.3261	0.3142	0.3710
Fuzzy en	0.3983	0.4789	0.3756	0.3225	0.3178	0.3180	0.3685
Rms	0.3919	0.3720	0.4007	0.5151	0.2872	0.3020	0.3781
Spectral skewness	0.4556	0.3257	0.3387	0.4457	0.2711	0.3033	0.3567
Spectral kurtosis	0.3386	0.2727	0.3110	0.3856	0.2435	0.2553	0.3011
Kurtosis	0.3734	0.0843	0.3260	0.3833	0.2422	0.3326	0.2903
Skewness	0.2248	0.1265	0.2261	0.2594	0.1712	0.1939	0.2003

PE are highest, followed by MSPE and RMS, which indicates a good degradation trend to reflect the bearing degradation process. Based on the CI evaluation, MMPE is found to be an effective feature for bearing degradation representation and is further used for HI construction and RUL prediction. The next step is to construct HI from the MMPE feature using Eq. (8) to indicate the overall degradation representation of the entire bearing. The original and cumulative effect of the MMPE feature for bearing 1_1 is shown in Fig. 8. It is clearly observable that the fluctuation in is more in the original feature and it does not reflect any trend, whereas constructed HI by considering absolute cumulative effect MMPE feature follows a decreasing monotonic trend and is further used for regression analysis.

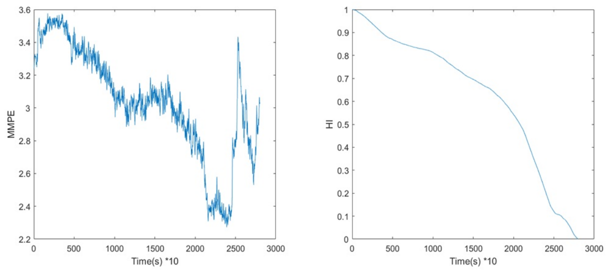

Fig. 8. The left side represents the original MMPE feature, whereas the right represents the HI constructed from the original MMPE for bearing 1_1.

4.1 Results and comparison of RUL prediction

The effectiveness of the proposed methodology is measured by comparing it with the other available methods in the literature.

The comparison of the score and mean absolute $ER\,\%$ is done, with the selected research work on bearing RUL prediction, ranging from 2018-2022, the details of which are being explained subsequently. The following RUL prediction works on bearings have been considered to substantiate the superiority of the proposed methodology: Hinchi and Tkiouat proposed a method by extracting the local features directly from the sensor in the convolution layer form using a neural network and then giving it as input to the LSTM model for RUL prediction of bearings [51]. Chen et al. proposed a deep learning-based data-driven approach with an attention mechanism for RUL prediction [52]. Zhang et al. proposed a hybrid deep learning network that can take both one-dimensional data and time-

Table 6. Rul prediction comparison with other methods.

Test bearings No.	Current time (s)	Actual RUL (s)	Predicted RUL by proposed method	Error%					
				Proposed method	Hinchi & Tkiouat [51]	Chen et al. [52]	Zhang et al. [53]	Wong et al. [54]	Xu et al. [55]
1_3	18,010	5730	5710	0.35	-0.35	1.05	2.27	5.06	-2.62
1_4	11,380	3390	3432	-1.24	5.60	20.35	5.6	23.30	17.40
1_5	23,010	1610	1720	-6.83	100.00	11.18	12.42	4.35	5.59
1_6	23,010	1460	1887	-29.23	28.08	34.93	10.96	0.68	3.42
1_7	15,010	7570	7467	1.36	-19.55	29.19	-22.46	-42.54	1.06
2_3	12,010	7530	7169	4.80	-20.19	57.24	0.99	17.40	26.96
2_4	6110	1390	1197	13.85	8.63	-1.44	5.76	12.23	-2.88
2_5	20,010	3090	2735	11.50	23.30	-0.65	25.89	-0.32	7.77
2_6	5710	1290	1232	4.51	58.91	-42.64	-10.85	-2.33	13.95
2_7	1710	580	571	1.62	5.17	8.62	1.72	8.62	-8.62
3_3	3510	820	868	-5.87	40.24	-1.22	-3.66	-3.66	3.66
$\overline{ER\%}$				-0.47	24.54	10.60	2.60	2.07	5.97
$\overline{\lvert ER\% \rvert}$				7.38	28.18	18.96	9.33	10.95	8.54
Score				0.81	0.43	0.57	0.64	0.67	0.69

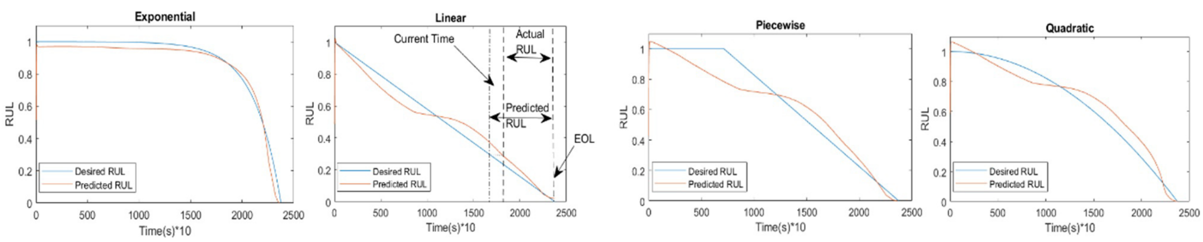

Fig. 9. Desired and predicted RUL for bearing 1_3 with four different target functions using HI constructed from C-MMPE.

frequency images as input for effective RUL prediction [53]. Wang et al. proposed a hybrid prognostic method by utilizing the sparse representation of degradation data and exponential degradation model for RUL estimation [54]. Xu et al. proposed a state degradation model and convolution autoencoder network to predict the RUL of bearings [55]. The comparative study of the results of the proposed and other methods are shown in Table 6. From the table, it is observed that the proposed method has the highest score of 0.81 and the lowest mean absolute $ER\%$ of 7.38. The highest score value indicates the strongest predictive capability and the better fitting of the model. The lowest mean absolute error percentage indicates the highest prediction accuracy between actual and predicted RUL. This illustrates that the proposed method can efficiently capture the C-MMPE based HI in time-series data, effectively predict the RUL with maximum accuracy compared to other methods and shows higher adherence to the requirements of practicability. In this way, work efficacy and the model ability are strengthened.

This work also compares the proposed method with other traditional features and labelling functions. Fig. 9 represents the plot between the desired and predicted RUL for all four labelling functions for bearing 1_3 using HI constructed from C-MMPE. The actual RUL is the time length between the current time and end of life (EOL), whereas the predicted RUL is the time length between the predicted time and EOL. It is clearly observed from Fig. 9, that the desired and predicted RUL for an exponential function is close to each other as compared to linear, piecewise, and quadratic functions. This indicates that the exponential labelling function provides better fitting and precisely predicts the RUL.

The score value and absolute ER% of each selected feature and labelling function are shown in Figs. 10 and 11, respectively. It's observed that the score value is maximum for the exponential labelling function for all selected features, showing its superiority among other labelling functions. Among features, the score value is maximum, i.e., 0.81 for HI constructed from C-MMPE followed by C-PE, C-MSPE, C-RMS, C-Skewness, and C-Kurtosis. Similarly, Fig. 11 shows that the absolute mean ER% is minimum, i.e., 7.38, for the exponential labelling function, and HI constructed from C-MMPE indicates the high accuracy between actual and predicted RUL. The results obtained from score value and absolute ER% suggest that the selected exponential labelling function and HI from C-MMPE are strong enough to predict the RUL precisely and accurately for rolling bearings.

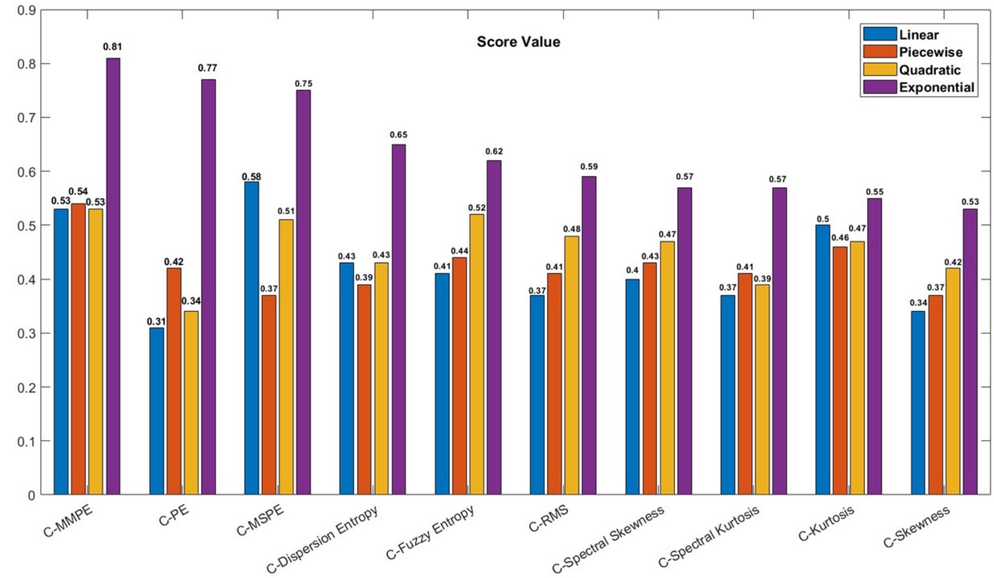

Fig. 10. Score value of HI constructed from all selected features and labelling functions.

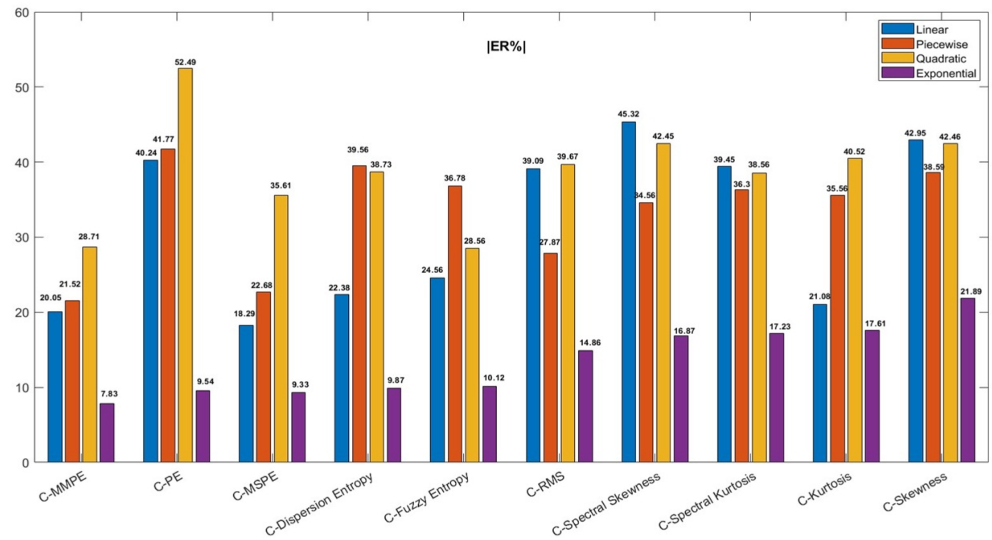

Fig. 11. Mean absolute error percentage of HI constructed from all selected features and labelling functions.

5. Conclusion

This paper presents a novel work for effective RUL prediction of the rolling bearing from a HI constructed from a single dominant feature C-MMPE and LSTM model. The performance of MMPE feature is evaluated from comprehensive indicator value and found to be higher than other features for both test and training datasets, indicating that the selected feature effectively represents the degradation behaviour of bearing. After the selection, HI is constructed by considering the cumulative effect of MMPE in decreasing patterns. The exponential output target function is subsequently defined to represent the virtual life or health degree of bearings. Finally, LSTM model is implemented for direct RUL prediction from the extracted features. The RUL prediction performance is measured by calculating the MAE and score value. The proposed method shows a low MAE value of 7.38 and a high score value of 0.81 as compared to other available methods in the literature, indicating the superiority of the model for RUL prediction. The same has been explained and displayed in a tabular format. The proposed method also shows its effectiveness with respect to other features such as RMS, skewness, kurtosis, PE, and MSPE and labelling functions such as linear, piecewise, and quadratic. However, in the future, there is a scope for reducing computational time due to the extensive feature extraction process. Hence, this work can be extended by developing novice intelligent feature extraction techniques for HI construction with the deep learning method.

Data availability

The data and material supporting this study's findings are openly available and provided by the FEMTO-ST Institute in PHM 2012 data repository.

Nomenclature

x_i : Original time series
$p(\pi)$: Relative frequency for each permutation 'π'
M : Embedding dimension
τ : Time lag
H_{PE} : Permutation entropy
H_{NPE} : Normalized permutation entropy
y_j^s : New time series sequence
f : Original feature
T : Time index
$\bar{f}$: Smoothing processing of the original feature
CI : Comprehensive indicator
t_n : Whole life duration of bearing
t_i : Current time
t_j : Initial degradation time of bearing
f_t : Forget gate
i_t : Input gate
o_t : Output gate
$\bar{C}_t$: Cell state
C_t : Update cell state
H_t : Output
$ER\%$: Error percentage
$\overline{ER\%}$: Mean error percentage
$\left|\overline{ER\%}\right|$: Mean absolute error percentage
A_i : Score for the i^{th} test bearing

Abbreviation

CG : Coarse graining
C-Prefix: Cumulative effect of features
HI : Health indicator
MAG : Moving average graining
MMPE : Modified multiscale permutation entropy
MSPE : Multi-scale permutation entropy
LSTM : Long short-term memory
PE : Permutation entropy
PHM : Prognostic health management
RMS : Root mean square
RUL : Remaining useful life

References

[1] Z. Liu and L. Zhang, A review of failure modes, condition monitoring and fault diagnosis methods for large-scale wind turbine bearings, *Measurement*, 149 (2020) 107002.
[2] A. Kumar, C. P. Gandhi, Y. Zhou, R. Kumar and J. Xiang, Improved deep convolution neural network (CNN) for the identification of defects in the centrifugal pump using acoustic images, *Applied Acoustics*, 167 (2020) 107399.
[3] X. Zhang, S. Wan, Y. He, X. Wang and L. Dou, Teager energy spectral kurtosis of wavelet packet transform and its application in locating the sound source of fault bearing of belt conveyor, *Measurement*, 173 (2021) 108367.
[4] Y. Wei, Y. Li, M. Xu and W. Huang, A review of early fault diagnosis approaches and their applications in rotating machinery, *Entropy*, 21 (4) (2019) 409.
[5] X. Kong and J. Yang, Remaining useful life prediction of rolling bearings based on RMS-MAVE and dynamic exponential regression model, *IEEE Access*, 7 (2019) 169705-169714.
[6] V. Atamuradov, K. Medjaher, P. Dersin, B. Lamoureux and N. Zerhouni, Prognostics and health management for maintenance practitioners-review, implementation and tools evaluation, *Int. J. Progn. Health Manag.*, 8 (3) (2017) 1-31.
[7] R. Llasag Rosero, C. Silva and B. Ribeiro, Remaining useful life estimation of cooling units via time-frequency health indicators with machine learning, *Aerospace*, 9 (6) (2022) 309.
[8] P. K. Sahu, R. N. Rai and T. C. A. Kumar, Grease contamination detection in the rolling element bearing using deep learning technique, *International Journal of Mechanical Engineering and Robotics Research*, 11 (4) (2022) 275-280.
[9] A. V Dube, L. S. Dhamande and P. G. Kulkarni, Vibration based condition assessment of rolling element bearings with localized defects, *Int. J. Sci. Technol. Res.*, 2 (4) (2013) 149-155.
[10] P. Arun, S. A. Lincon and N. Prabhakaran, Detection and characterization of bearing faults from the frequency domain features of vibration, *IETE J. Res.*, 64 (5) (2018) 634-647.
[11] X. Zhang, Y. Liang and J. Zhou, A novel bearing fault diagnosis model integrated permutation entropy, ensemble empirical mode decomposition and optimized SVM, *Measurement*, 69 (2015) 164-179.
[12] P. K. Sahu and R. N. Rai, Fault diagnosis of rolling bearing based on an improved denoising technique using complete ensemble empirical mode decomposition and adaptive thresholding method, *Journal of Vibration Engineering & Technologies*, 11 (2023) 513-535.
[13] Y. Li, X. Wang, Z. Liu, X. Liang and S. Si, The entropy algorithm and its variants in the fault diagnosis of rotating machinery: A review, *IEEE Access*, 6 (2018) 66723-66741.
[14] C. E. Shannon, A mathematical theory of communication, *The Bell System Technical Journal*, 27 (3) (1948) 379-423.
[15] S. Pincus, Approximate entropy (ApEn) as a complexity measure, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 5 (1) (1995) 110-117.
[16] J. S. Richman and J. R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, *American Journal of Physiology-Heart and Circulatory Physiology*, 278 (6) (2000) H2039-H2049.
[17] C. Bandt and B. Pompe, Permutation entropy: A natural complexity measure for time series, *Phys. Rev. Lett.*, 88 (17) (2002) 174102.
[18] R. Yan, Y. Liu and R. X. Gao, Permutation entropy: A nonlinear statistical measure for status characterization of rotary machines, *Mech. Syst. Signal Process*, 29 (2012) 474-484.

[19] D.-Y. Lee and Y.-S. Choi, Multiscale distribution entropy analysis of short-term heart rate variability, *Entropy*, 20 (12) (2018) 952.

[20] Q. Xue et al., Feature extraction using hierarchical dispersion entropy for rolling bearing fault diagnosis, *IEEE Trans. Instrum. Meas.*, 70 (2021) 1-11.

[21] J. Jiao, J. Yue and D. Pei, Feature enhancement method of rolling bearing based on K-adaptive VMD and RBF-fuzzy entropy, *Entropy*, 24 (2) (2022) 197.

[22] H. Zhang and S. He, Analysis and comparison of permutation entropy, approximate entropy and sample entropy, *2018 International Symposium on Computer, Consumer and Control (IS3C)*, Taichung, Taiwan (2018) 209-212.

[23] M. Costa, A. L. Goldberger and C.-K. Peng, Multiscale entropy analysis of complex physiologic time series, *Phys. Rev. Lett.*, 89 (6) (2002) 068102.

[24] H. Ren, W. Liu, M. Shan and X. Wang, A new wind turbine health condition monitoring method based on VMD-MPE and feature-based transfer learning, *Measurement*, 148 (2019) 106906.

[25] Y. Li, M. Xu, Y. Wei and W. Huang, A new rolling bearing fault diagnosis method based on multiscale permutation entropy and improved support vector machine based binary tree, *Measurement*, 77 (2016) 80-94.

[26] S.-D. Wu, C.-W. Wu, K.-Y. Lee and S.-G. Lin, Modified multiscale entropy for short-term time series analysis, *Physica A: Statistical Mechanics and its Applications*, 392 (23) (2013) 5865-5873.

[27] C. Su, Z. Liang, X. Li, D. Li, Y. Li and M. Ursino, A comparison of multiscale permutation entropy measures in on-line depth of anesthesia monitoring, *PLoS One*, 11 (10) (2016) e0164104.

[28] B. Zhang, L. Zhang and J. Xu, Degradation feature selection for remaining useful life prediction of rolling element bearings, *Qual. Reliab. Eng. Int.*, 32 (2) (2016) 547-554.

[29] L. Duan, F. Zhao, J. Wang, N. Wang and J. Zhang, An integrated cumulative transformation and feature fusion approach for bearing degradation prognostics, *Shock and Vibration*, 2018 (2018) 9067184.

[30] P. K. Sahu and R. N. Rai, Degradation monitoring and RUL prediction of rolling element bearing using proposed C-MMPE feature, *2022 11th International Conference on Power Science and Engineering (ICPSE)*, Eskisehir, Turkey (2022) 54-60.

[31] D. Chen, Y. Qin, Y. Wang and J. Zhou, Health indicator construction by quadratic function-based deep convolutional autoencoder and its application into bearing RUL prediction, *ISA Trans.*, 114 (2021) 44-56.

[32] C. Wang, W. Jiang, X. Yang and S. Zhang, RUL prediction of rolling bearings based on a DCAE and CNN, *Applied Sciences*, 11 (23) (2021) 11516.

[33] X.-S. Si, W. Wang, C.-H. Hu and D.-H. Zhou, Remaining useful life estimation–a review on the statistical data driven approaches, *Eur. J. Oper. Res.*, 213 (1) (2011) 1-14.

[34] J. Zhu, N. Chen and W. Peng, Estimation of bearing remaining useful life based on multiscale convolutional neural network, *IEEE Transactions on Industrial Electronics*, 66 (4) (2018) 3208-3216.

[35] R. Liu, B. Yang and A. G. Hauptmann, Simultaneous bearing fault recognition and remaining useful life prediction using joint-loss convolutional neural network, *IEEE Trans. Industr. Inform.*, 16 (1) (2019) 87-96.

[36] C. Sun, M. Ma, Z. Zhao, S. Tian, R. Yan and X. Chen, Deep transfer learning based on sparse autoencoder for remaining useful life prediction of tool in manufacturing, *IEEE Trans. Industr. Inform.*, 15 (4) (2018) 2416-2425.

[37] J. Deutsch and D. He, Using deep learning-based approach to predict remaining useful life of rotating components, *IEEE Trans. Syst. Man. Cybern. Syst.*, 48 (1) (2017) 11-20.

[38] J. Lei, C. Liu and D. Jiang, Fault diagnosis of wind turbine based on long short-term memory networks, *Renew. Energy*, 133 (2019) 422-432.

[39] S. Zheng, K. Ristovski, A. Farahat and C. Gupta, Long short-term memory network for remaining useful life estimation, *2017 IEEE International Conference on Prognostics and Health Management (ICPHM)*, Dallas, TX, USA (2017) 88-95.

[40] C.-S. Hsu and J.-R. Jiang, Remaining useful life estimation using long short-term memory deep learning, *2018 IEEE International Conference on Applied System Invention (ICASI)*, Chiha, Japan (2018) 58-61.

[41] J. Xia, Y. Feng, C. Lu, C. Fei and X. Xue, LSTM-based multi-layer self-attention method for remaining useful life estimation of mechanical systems, *Eng. Fail. Anal.*, 125 (2021) 105385.

[42] W. Mao, J. He, J. Tang and Y. Li, Predicting remaining useful life of rolling bearings based on deep feature representation and long short-term memory neural network, *Advances in Mechanical Engineering*, 10 (12) (2018).

[43] M. S. Rathore and S. P. Harsha, Rolling bearing prognostic analysis for domain adaptation under different operating conditions, *Eng. Fail Anal.*, 139 (2022) 106414.

[44] M. Rostaghi and H. Azami, Dispersion entropy: A measure for time-series analysis, *IEEE Signal Process Lett.*, 23 (5) (2016) 610-614.

[45] W. Chen, Z. Wang, H. Xie and W. Yu, Characterization of surface EMG signal based on fuzzy entropy, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 15 (2) (2007) 266-272.

[46] K. Javed, R. Gouriveau, N. Zerhouni and P. Nectoux, Enabling health monitoring approach based on vibration data for accurate prognostics, *IEEE Transactions on Industrial Electronics*, 62 (1) (2014) 647-656.

[47] Y. Yu, X. Si, C. Hu and J. Zhang, A review of recurrent neural networks: LSTM cells and network architectures, *Neural Comput.*, 31 (7) (2019) 1235-1270.

[48] F. A. Gers, J. Schmidhuber and F. Cummins, Learning to forget: continual prediction with LSTM, *Neural Comput.*, 12 (10) (2000) 2451-2471.

[49] J. Bae and Z. Xi, Learning of physical health timestep using the LSTM network for remaining useful life estimation, *Reliab. Eng. Syst. Saf.*, 226 (2022) 108717.

[50] P. Nectoux et al., PRONOSTIA: An experimental platform for bearings accelerated degradation tests, *IEEE International*

Conference on Prognostics and Health Management, Denver, USA (2012) 1-8.

[51] A. Z. Hinchi and M. Tkiouat, Rolling element bearing remaining useful life estimation based on a convolutional long-short-term memory network, *Procedia Comput. Sci*., 127 (2018) 123-132.

[52] Y. Chen, G. Peng, Z. Zhu and S. Li, A novel deep learning method based on attention mechanism for bearing remaining useful life prediction, *Appl. Soft Comput.*, 86 (2020) 105919.

[53] G. Zhang, W. Liang, B. She and F. Tian, Rotating machinery remaining useful life prediction scheme using deep-learning-based health indicator and a new rvm, *Shock and Vibration*, 2021 (2021) doi: 10.1155/2021/8815241.

[54] B. Wang, Y. Lei, N. Li and N. Li, A hybrid prognostics approach for estimating remaining useful life of rolling element bearings, *IEEE Trans. Reliab.*, 69 (1) (2018) 401-412.

[55] W. Xu, Q. Jiang, Y. Shen, F. Xu and Q. Zhu, RUL prediction for rolling bearings based on convolutional autoencoder and status degradation model, *Appl. Soft. Comput.*, 130 (2022) 109686.

Prashant Kumar Sahu is presently working as a Research Scholar under the guidance of Dr. Rajiv Nandan Rai at Subir Chowdhury School of Quality and Reliability, Indian Institute of Technology Kharagpur, West Bengal, India. He received his B.E degree in Mechanical Engineering from Sathyabama University, Chennai in 2014 and MTech in Maintenance Engineering and Tribology from IIT (ISM) Dhanbad in 2017. His research interests include prognostic health management for rotating machinery.

Rajiv Nandan Rai, a Ph.D. in Mechanical Engineering from IIT Delhi, and presently a faculty at IIT Kharagpur, is an accomplished engineer with over 29 years of experience in the field of Reliability, Quality and Maintenance Engineering. He has authored more than 30 publications (SCI and Scopus indexed) including 02 book and 05 book chapters. He has also completed a number of sponsored Projects. His research interests include Reliability analysis of repairable systems, maintenance engineering, quality management and engineering, machine diagnostics and prognostics.