

Article

Intelligent Online Monitoring of Rolling Bearing: Diagnosis and Prognosis

Hassane Hotait ^{1,*}, Xavier Chiementin ¹ and Lanto Rasolofondraibe ²

¹ ITheMM, Institute of Thermics, Mechanics and Material, University of Reims Champagne-Ardenne, Moulin de la Housse, 51687 Reims, France; xavier.chiementin@univ-reims.fr

² CReSTIC, University of Reims Champagne-Ardenne, Moulin de la Housse, 51687 Reims, France; lanto.rasolofondraibe@univ-reims.fr

* Correspondence: hassane.hotait@etudiant.univ-reims.fr; Tel.: +33-767-206-074

Abstract: This paper suggests a new method to predict the Remaining Useful Life (RUL) of rolling bearings based on Long Short Term Memory (LSTM), in order to obtain the degradation condition of the rolling bearings and realize the predictive maintenance. The approach is divided into three parts: the first part is the clustering to detect the damage state by the density-based spatial clustering of applications with noise. The second one is the health indicator construction which could give a better reflection of the bearing degradation tendency and is selected as the input for the prediction model. In the third part of the RUL prediction, the LSTM approach is employed to improve the accuracy of the prediction. The rationale of this work is to combine the two methods—the density-based spatial clustering of applications with noise and LSTM—to identify the abnormal state in rolling bearings, then estimate the RUL. The suggested method is confirmed by experimental data of bearing life cycle, and the RUL prediction results of the model LSTM are compared with the nonlinear au-regressive model with exogenous input model. In addition, the constructed health indicator is compared with the spectral kurtosis feature. The results demonstrated that the suggested method is more appropriate than the nonlinear au-regressive model with exogenous input model for the prediction of bearing RUL.

Keywords: bearings; vibration signal; DBSCAN; health indicator; LSTM; RUL

Citation: Hotait, H.; Chiementin, X.; Rasolofondraibe, L. Intelligent Online Monitoring of Rolling Bearing: Diagnosis and Prognosis. *Entropy* **2021**, *23*, 791. <https://doi.org/10.3390/e23070791>

Academic Editors: Claude Delpha and Demba Diallo

Received: 6 May 2021

Accepted: 31 May 2021

Published: 22 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Rolling bearings, which is widely used in rotating machinery, are among the main parts of mechanical equipment. As the equipment runs, their performance will deteriorate and lead to failure. Consequently, early fault diagnosis and prediction of bearing life and safety are of high importance for predictive maintenance and the industrial reliability of mechanical equipment [1].

The vibration signals which are generated by rotating machines and collected can reflect the different processing conditions of modern industrial equipment. Industrial equipment is becoming more and more sophisticated and sensitive as technology evolves. For this reason, availability, reliability, and reduction of downtime and repairs are important. In that matter, condition monitoring has been determined as an effective means of increasing safety, health, and optimal equipment performance. Condition monitoring means the damage assessment and maintenance of machines without disturbing their operation and is performed based on information obtained from the condition of the equipment [2]. The vibration signals are used to provide real-time monitoring of the different states of bearings. The presence of intense vibrations in the rotating machinery usually implies the occurrence of faults in bearings. The vibration method is the base to start the data classification, and then detect the abnormal state of rolling bearings [3]. The density-

based spatial clustering of applications with noise (DBSCAN) is an efficient method recently used to classify data [4]. The DBSCAN algorithm uses the signal density to separate the dense region from the noise region. Several analyses of bearing fault diagnostics and condition monitoring have been done as well as studying the prognostics of the defect in rolling bearings [5]. These predictions can reduce machine breakdowns, thereby minimizing the costs and support maintenance scheduling thus prolonging the life of the equipment.

Generally, data-driven prognostic frameworks comprise four stages: data acquisition, health indicator (HI) construction, degradation modeling and Remaining Useful Life (RUL) prediction [6]. Spectral kurtosis is used as a health indicator for wind turbine high-speed shaft bearings health prognosis analysis [7] as well as for the damage level assessment in bearing diagnosis [8]. Concerning the data acquisition, the diversity and the amount of sensor data collected in real-time must be enough to completely characterize the degradation performance of each component. For rotating machinery components, vibration signals are collected regularly, and time domain and frequency domain analyses are used for HI extraction [9]. One of the main queries in the prognostic of bearings is how to construct and evaluate health indicators from accessible features, which can represent the degradation states. The health indicator can be extracted from the time domain, frequency domain and scale domain [10].

For construction, the health indicator must reduce the number of features by using the Kernel Principal Component Analysis (KPCA) method [11] which is one of the most importantly used techniques. After the extraction of the health indicator, the mission is to choose the best model to predict the remaining useful health of the model. With the elaboration of sensor technologies, many data-driven methods along with statistical models have been developed to define product failures [12]. Particularly, the remaining useful life (RUL) of rotating machinery is described as the period between the present monitoring time and the failure time when the degradation signal surpasses a failure threshold level [13]. RUL prediction based on degradation signals has evolved as a crucial technology that gives failure information for health management and condition-based maintenance [14]. The data-driven RUL prediction methods that are widely used comprise regression models, Bayesian reasoning [15], Gaussian mixture models [16], and other statistical analysis methods, in addition to artificial intelligence methods such as fuzzy decision tree [17], artificial neural network (ANN) [18], and hidden Markov model [19]. The deep learning method is a hopeful tool to reach real-time bearing fault diagnosis, because it can provide a stable hierarchical feature representation from the raw bearing vibration signal and then investigate the bearing status online according to the real-time collected vibration signal. The importance of deep learning consists of its high potential to bypass false features caused by ambient noise and fluctuations in working conditions [20].

The purpose of RUL prediction is to obtain the deterioration trend of bearings at the present moment. In the part of RUL prediction, some samples are inserted into a Long Short Term Memory (LSTM) network in clusters as training sets to achieve model training and network parameter arrangement. The constructed model is evaluated with test sets, and the RUL prediction rates of these sets are attained. ANNs with various configurations has been utilized for short-term wind forecasting, where estimation beyond the last observations collected in the training data are only devoted to a few time steps [21]. An ANN (consisting of the input layer, hidden layer and output layer) is named shallow if it uses one hidden layer and deep if it utilizes more than one [22]. The recurrent neural network (RNN), is an inner structure of deep neural networks, as it makes the repair of complicated nonlinearities in the data feasible. It contains memory blocks with the possibility for reminding the information at each time from the preceding samples [23]. The LSTM is one of the well-known kinds of RNN for processing time series that can hold information for a long duration during the learning process [24]. Billings developed the Nonlinear Auto-Regressive with Exogenous Inputs (NARX) model as a new description of a large class of discrete, nonlinear systems [25]. Many systems have been examined by utilizing the

NARX model [26]. It is another kind of RNN in which prediction of wind speed can be achieved by a procedure with additional meteorological time series data like air temperature and wind orientation [27]. Also, the NARX model can be used in other industrial structures such as modeling the major horizontal axis of wind turbines [28].

This study tests the predictability of rolling bearings by utilizing two different deep learning data-driven methods, the LSTM network model and the NARX that is also employed to compare and prove the effectiveness of the proposed method.

The achievability of the proposed process has been validated with experimental data sets. The following is a compendium of the major achievements of this paper:

- (1) Data classification to detect an abnormal state in rolling bearings.
- (2) Building of a special HI based on KPCA dimension reduction. This can efficaciously illustrate the degradation of rolling bearing operation.
- (3) Presentation of a deep prognosis network—LSTM—which ameliorates the accuracy of the HI and RUL estimations significantly.
- (4) Proof of the potential of the proposed methodologies based on experimental rolling bearing datasets. The results show that LSTM accomplished better execution than the NARX prognosis approach.

This paper is organized as follows: Section 1 is an introduction presenting a general description of the classification and reduction of dimensions to construct the health indicators, and finally the use of the LSTM neural network model and RUL. Section 2 describes the DBSCAN method used to classify and detect the degradation states. Section 3 presents the LSTM method to predict the remaining useful life of the studied equipment. Section 4 discusses the methodology proposed to detect the degradation state and predict the RUL. Section 5 presents the experimental verification of the suggested method. Then, Section 6 presents a comparison of the proposed methodology with NRAX. Lastly, Section 7 puts forth the conclusions of the paper.

2. DBSCAN Method

Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is a well-known density-based clustering algorithm, first suggested by Ester [29], that is explained as a density-based clustering non-parametric algorithm in-favor-of a set of points in space. It is based on linking points that are closely stuffed together in high-density zones, thereby classifying as outliers the points that are isolated in low-density regions. It is considered one of the representative clustering algorithms and remarkably widely cited in the scientific literature. It implies the number of clusters based on the data rather than using it as requirement parameter. It can also detect clusters of arbitrary shape. The ϵ -neighborhood is essential for DBSCAN to estimate local densities, so the algorithm has two parameters: ϵ , which is the radius of a neighborhood near a data point p , in addition to MinPts , that is the minimal number of data points in a neighborhood needed to reveal a cluster.

Utilizing these two parameters, DBSCAN layers the data points in three categories:

Core Points are the foundations for clusters based on the density. The identical ϵ used to calculate the neighborhood for every point so the core points are data points that suited a minimal density requirement. A data point p is a core point if neighborhood (p, ϵ) [ϵ -neighborhood of p] includes at least MinPts ; $|\text{neighborhood}(p, \epsilon)| \geq \text{MinPts}$.

Border Points are the points in clusters other than core points. A data point q is a border point if neighborhood (q, ϵ) comprises less than MinPts data points, however q is density-reachable of several core point p .

Outliers are points that are not referred to any cluster and maybe considered noisy points. The data point o is considered as outlier if it is none of the two, core point or border point. In other words, it is the “other” class.

The DBSCAN algorithm pseudocode can be expressed as the Algorithm 1:

Algorithm 1. DBSCAN algorithm pseudocode.

1. Entries: DBSCAN ($M, \epsilon, \text{MinPts}$)
 2. Outputs: Class sets
 3. Initialization arbitrarily choose a point p
 4. As long as (there are still untreated points)
 5. Find all density-reachable points of p using ϵ and MinPts
 6. If (p is a point of type center) a cluster is formed
 7. If not
 8. If (p is border type point and no point is reachable by density of p), the point p is considered as a noise and DBSCAN goes to the next point in the database.
 9. End
-

3. Long Short-Term Memory

Long short-term memory (LSTM) is a type of artificial recurrent neural network (RNN) [30] used in the field of deep learning. LSTM networks are used for classifying, processing, and making predictions based on time series data in the context of the occurrence of unknown duration delay between important events in a time series.

LSTMs were developed to treat the vanishing gradient issues that can be encountered when training traditional RNNs. They are superior to RNNs, hidden Markov models, and other sequence learning methods in several applications [31]. The advantage consists predominantly in the LSTMs' ability to hold long-term memory, whereas due to the short-term memory, the typical RNN will only be capable of utilizing the contextual information from the last data which is not helpful at all [32]. They also have a multifactorial structure. At every time step, the LSTM cell considers three different factors: the current input data, the short-term memory (hidden state), and lastly the long-term memory (cell state). The cell then uses gates to analyze the information to be kept at each time step before reaching the long-term and short-term information to the next cell. These gates are called the Input Gate, the Forget Gate, and the Output Gate [33]. These gates are defined as follows:

Input Gate. This chooses what new information will be kept in the long-term memory. It only handles the information from the present input and the short-term memory from the preceding time step, so it has to extract the non-useful information from these variables.

Forget Gate. It settles which information from the long-term memory should be preserved. This is achieved by multiplying the incoming long-term memory by a forget vector produced by the present input and incoming short-term memory.

Output Gate. This will use the present input, the preceding short-term memory, and the recently computed long-term memory to obtain the new short-term memory, which will be moved on to the cell in the following time step. The output of the present time step can also be displayed from the hidden state.

A typical LSTM module, also known as a repeating module, has four neural network layers linked to each other uniquely as shown in Figure 1. The module has three-gate activation functions σ and two output activation functions ϕ as described in Figure 1. This is usually similar to replacing the addition operation (+) by a Hadamard product (\odot) in the standard Elman-RNN equation. The concatenation operation is represented by the symbol (\circ) bullet. The network can indicate the amount of previous information to flow. It is controlled through the first layer (σ), given by Equation (1) [34].

Given that X_t the input gate activator vector at time t , o_t the forget gate activator vector at time t , c_t the outcome of cell vector at time t , h_t the outcome of layer vector at time t , W_x , W_i , and W_o the input kernels for the respective gates, U_f , U_i , and U_o , the recurrent kernels for the respective gates, b_f , b_i , b_o , and

b_c the biases, σ the logistic sigmoid function, \odot the matrix product, and \tanh the hyperbolic tangent activation function.

$$x_t = \sigma(\mathbf{W}_x x_t + \mathbf{U}_x h_{t-1} + b_f) \tag{1}$$

$$i_t = \sigma(\mathbf{W}_i x_t + \mathbf{U}_i h_{t-1} + b_i) \tag{2}$$

$$o_t = \sigma(\mathbf{W}_o x_t + \mathbf{U}_o h_{t-1} + b_o) \tag{3}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(\mathbf{W}_c x_t + \mathbf{U}_c h_{t-1} + b_c) \tag{4}$$

$$h_t = o_t \odot \tanh(c_t) \tag{5}$$

The current information to be stored in the cell state is generated using two network layers. A sigmoid layer (σ) that establishes values to update (i_t) (Equation (2)) and tanh layer ϕ that details a vector of new possible values (o_t) as shown in Equation (3). The fusion of both is then added in the state. Lastly, the cell state is modified using Equation (5) [35].

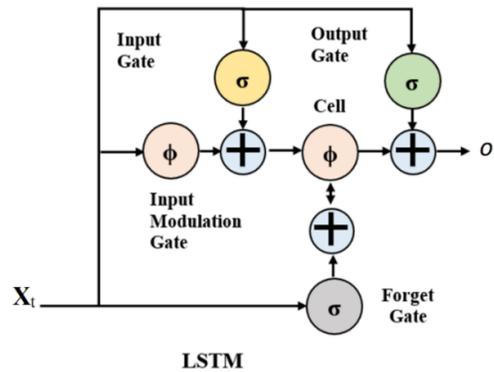


Figure 1. The LSTM module, where X_t , i_t , o_t are input, and output gates, respectively.

4. Details of the Proposed Methodology

The proposed methodology consists of online-integrated chains that are shown in Figures 2 and 3, which make it robust. It starts by determining the healthy state of rolling bearings and then stratifying the data into healthy and faulty. After detection of the defect, the prognostics neural network is constructed, that aims to calculate the remaining useful life. The objective of this methodology is to detect the defect by the DBSCAN method, then calculate the remaining useful life by the LSTM method to predict the equipment degradation state evolution.

Figure 1 reflects the proposed approach, which is divided into three parts: the first part is the clustering to detect the damage state, the second one is the health indicator construction, the third part of the RUL and HI prediction.

4.1. The Proposed Framework

The chain begins with the data acquisition by collecting signals, then extracting features. The collected features are described in Section 4.2 Feature Extraction Methods, where they are separated into a time-domain. After their extraction, these features are normalized by the z-score method [36] which is described by Equation (6). The KPCA method was used to reduce the number of features and show them in 3D. The DBSCAN parameters Minpts and Epsilon are defined in this stage. The calculation of the Minpts value depends on the initial matrix T which is equal to the first two successive moments (every successive moments contain eight signals) and the Minpts value is equal to $(\frac{1}{2}T)$.

The epsilon is calculated by the nearest neighbor method for the initial matrix (Equation (7)). The estimation of epsilon value is realized by generating a k-distance graph for

the input data MF . For every point in MF , the distance to the k -th nearest point is calculated, then arranged points at this distance are plotted. The graph includes a knee. The distance connected to the knee is mostly a good choice for epsilon since it is the region where points start appending off into outlier (noise) region [37].

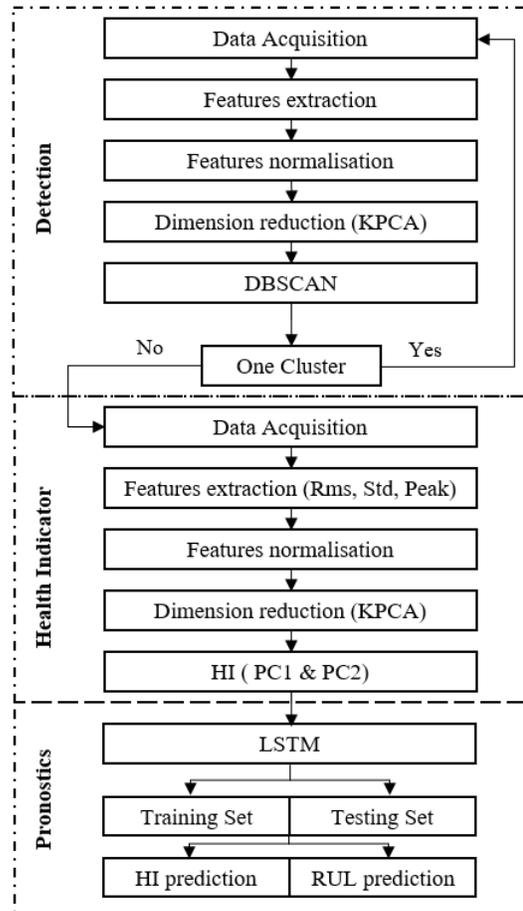


Figure 2. Flowchart of the proposed methodology.

Prior to plotting the k -distance graph, the Minpts that are the short pairwise distances for examinations in MF have to be sited in ascending order. The features matrix MF , is normalized $[MF]_i^{norm}$ (Equation (6)). The objective of normalization is to transform the calculated values to be on a similar scale:

$$[MF]_i^{norm} = \frac{MF_i - \overline{MF}_i}{std(MF_i)} \tag{6}$$

ε corresponds to the maximum distance between the center of the class, c_h and the MinPtsth neighbor. The outcoming class is so-called healthy class, well known C_h , with center c_h . This class represents a reference state:

$$\varepsilon = \text{distance}(c_h, \text{MinPts}^{\text{th}} \text{ neighbor}) \tag{7}$$

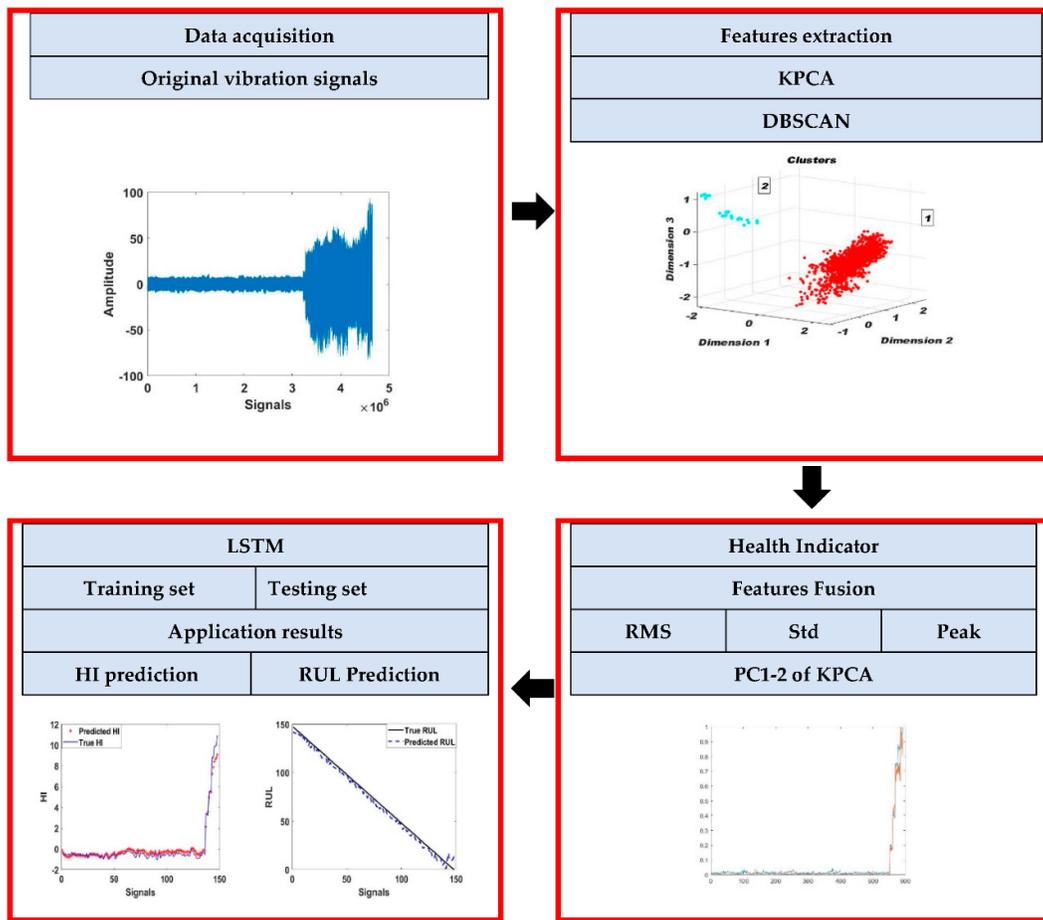


Figure 3. The procedure of the proposed methodology.

The DBSCAN objective is to detect the degradation state. When two or more classes are detected the first one is the normal state, the second and the third are the degraded ones. This chain returns to the acquisition phase under the condition of not detecting the second class. This classification ends with the detection of the second class, which is the degraded class [38].

The prognostic’s objective is to show the evolution of the degraded state of the equipment and calculate the RUL. The health indicator is divided into trained and tested data [39]. After that, the application of the regression model will be done by the LSTM method to predict the remaining useful life.

4.2. Health Indicator

In the field of fault diagnosis, the time domain features are an essential index to detect the health condition of mechanical equipment, three pertinent time domain features are combined and fused with the KPCA method to construct the health indicator, which can be defined as follows in Table 1.

Table 1. Health indicator features.

$$\text{Root mean square: } RMS = \left(\frac{1}{N} \sum_{i=1}^N (x_i^2) \right)^{\frac{1}{2}} \quad (8)$$

$$\text{Standard deviation: } (Std) = \left(\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \right)^{\frac{1}{2}} \quad (9)$$

$$\text{Peak to Peak: } x_{PEAK} = \max(x_i) - \min(x_i) \quad (10)$$

Algorithm 2: HI construction algorithm.

Algorithm 2. HI construction algorithm.

Input: raw life-cycle data; combine (RMS, Std, Peak) features;
Sensitive health features selection.

1. Construct a feature space with 3 features referred to in Table1 by time-domain, from raw life-cycle data.
2. Reduction of the dimension of 3 features by KPCA.
3. Minimize the weights of each input-sensitive health feature by KPCA.
4. The first Two principal component (1-2) represents the maximum variance direction in the data.
5. Find the sensitive health features by the predetermined termination.
6. Construct the optimal one-dimensional HI.

Output: HI; sensitive health features.

4.3. Feature Extraction Methods

The multi-domain feature set can completely represent the bearing fault feature state which can provide an efficient diagnosis for different rolling bearing faults occurring under varying speed and load or anonymous speed conditions. The existence of defects in-machinery components can be observed in the raw acceleration signals. To get a better comprehension of the vibration signals we attempt to extract the time domain, frequency domain, and time-frequency domain features [36].

Time-domain features of vibration signals have proved to be useful to present the machinery condition. The time-domain features were used to detect bearing damages such as kurtosis, standard deviation, skewness, root mean square efficiency for detection, and localized bearing default. In addition to the peak-to-peak value, impulse factor, Tikhat, Talaf, crest factor, that can detect the changes in the signal when a defect occurs. These classical time-domain statistical features were applied to represent the bearing condition [40].

The frequency-domain is based on the spectral analysis technique that can convert time-domain vibration signals into discrete frequency components using a fast Fourier transform (FFT) [41]. The signal spectrum contains rich condition information. The frequency root mean square f_{rms} , and the frequency root mean square brut: f_{rmsb} , the frequency center: f_c , standard deviation frequency brut: f_{stdb} , the power envelope: PW , the mean frequency f_{rmsf} are used in the rolling bearing to detect the defect [42].

The time-scale analysis methods decompose the extracted signals into a set of scale components that contain several fault features. The wavelet process has an advantage over the conventional Fourier transform in the condition of the analysis of discontinuous and short-impulse signals [43]. Wavelet is a well-known signal processing technique used to examine the frequency composition of the signal [44]. Therefore, technical description will be omitted. Instead, the two new features extracted will be introduced: WRMS, which is the effective value of the frequencies of the signal's wavelets spectrum, and PCWT that represents the average value of the envelope amplitudes [45].

4.4. Reduction of Dimensions

The reduction of dimensions is a step used to reduce the number of features and easily plot the observed data in 3D or 2D. There is often a correlation between the feature dimensions and the fault detection. Thereby, it is very significant to extract the critical fault features that reproduce the fault and reduce the dimension data. KPCA is a simple dimension reduction method which is an extension of PCA. It is an efficient tool for multivariable and nonlinear data, and its major functions are that it plots the indigenous spatial data into high dimensional space by kernel function, converts the original nonlinear issue to linearization one, and utilizes PCA to minimize the dimensions [44].

4.5. Evaluation of Remaining Useful Life (RUL)

The performance of the RUL estimation procedure has been tested extensively to evaluate the RUL in the fitted LSTM method. According to the criteria adopted in the methodology (Figure 2), three evaluation criteria are taken into consideration to measure the prediction performance of different approaches:

4.5.1. Root Mean Square Error (RMSE)

Root mean square error is often used to measure of the differences within predicted values by a model or observed and estimator values. The RMSE represents the square root of the second sampling time of the differences between predicted and observed values or the root mean square of these differences [46].

The RMSE of prognosticate values y_i for times t of a regression's count on variable x_i with observed variables over n are the signals number. It is calculated for T different predictions and is equal to the square root of the mean of the squares of the deviations:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - x_i)^2}{T}} \quad (11)$$

4.5.2. Mean Absolute Error (MAE)

Mean absolute error (MAE) is the calculation of errors among paired observations noticing the same occurrence. Patterns of Y versus X include contrasts of predicted against observed, initial time versus subsequent time, and unitechnique versus an alternative technique of measurement [47]. MAE is calculated as:

$$\text{MAE} = \frac{\sum_{i=1}^n |y_i - x_i|}{n} = \frac{\sum_{i=1}^n |e_i|}{n} \quad (12)$$

The mean absolute error is calculated as follows: $|e_i| = |y_i - x_i|$ where y_i is the prediction value and x_i is the true value. It utilizes the same scale as the existent in the measured data and is a familiar measure of forecast errors in time series analysis.

5. Experimental Validation

5.1. Test Bench

The tests were performed on a group of identical single-row thrust bearings. The objective was to follow the development of spalling from micro to fatigue state. These tests were performed on a fatigue module of a test bench. A constant axial load of 3000 daN, was applied to the bearing with the rotation speed remained constant during the tests, i.e., 1800 rpm. The coolant flow was also constant. Two type DJB3208 and DJB3209 piezoelectric accelerometers were placed as close as possible to the bearing in two different directions, axial and radial. The used data is the data of the piezoelectric sensor positioned radially on the bearing is taken as the best measuring point. The OROS OR34 acquisition system recorded 8192 points of the vibratory signal in a frequency range of 20 kHz to have a significant number of cycles.

The test bench shown in Figure 4, is essentially composed of an electric motor (1), that turns a spindle (2) on which one of the bearing rings is mounted. A shaft (3) transmits the axial load to the thrust bearing from a hydraulic pump (4), and a continuously operating lubrication circuit (5) for cooling. This test bench is connected to a control and data acquisition system. The ball bearings used are of type FAG 51207 CZECH/ATK. This bearing allows easy disassembly and immediate visual inspection of defects.



Figure 4. Module of the thrust bearing fatigue test bench.

5.2. Procedure of the Experiments

During the tests, the temperatures of the main shaft bearing, the lubricating oil, and the ambient air were continuously monitored. As shown in Table 2, three bearings were used and subjected to a fatigue phenomenon until the bearing broke, figure 5. The health indicator was built from the first two principal components which covered 99.1% of the original data but the first component represented less with 97.9%.

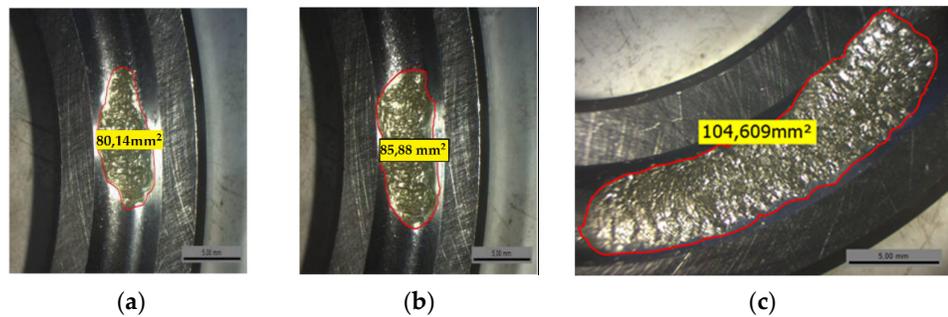


Figure 5. The spalling bearing of the three tested rolling bearing for the bearing number 1 (a), (b) for bearing number 2 and (c) for the bearing number 3.

The cases 1–3 were applied on the proposed method LSTM to get the predicted health indicator and RUL. Knowing that cases 1 and 2 were the training cases of the bearing 1 with the testing of the bearings 2 and 3, respectively. As well, case 3 was the training of the bearing 2 with the testing of the bearing 3, with 25% of the training data being tested in each case.

Table 2. Summary of fatigue tests on bearings.

Bearing Number	Signal Number	Spalling Size (mm ²)
1	592	80.14
2	592	85.88
3	592	104.6

5.3. Results of the Experiments

Three different conditions were considered throughout the experiment, operating with different spalling sizes: the first one with 80.14 mm², the second one with 85.88 mm² and the third one with 104.6 mm². The measured vibration signals of bearing numbers 1–3 are exhibited in Figure 6a–c. During the tests, every bearing was naturally degraded. As a result, every test bearing collected a different vibration signal pattern.

As shown in Figure 4, the entire lifetime vibration signals of three test bearings were demonstrated. The signal amplitudes of bearing 1 displayed a fast increase near the end of life which designated a sudden degradation. The signal amplitudes of bearings 2 and 3 had a trend of progressive increase. This showed that degradation began to change slightly at first and then in a severe way.

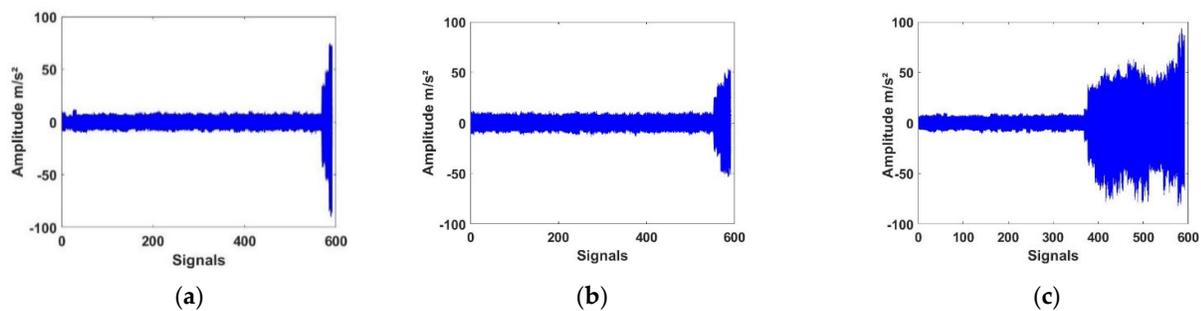


Figure 6. The time-domain waveform of bearing, (a) number 1, (b) number 2, and (c) number 3.

To determine the unique optimal value Eps, rate determination for single-level density was obtained by calculating the slope between points. Figure 7 depicts the result of a plot that has been sorted in ascending values. Detection of Eps value was done by calculating the slope of the lines of four Minpts values. The slopes of the lines were located for the three bearings at the points of 0.170, 0.229 and 0.272, which are the optimal values Eps for bearings 1, 2 and 3, respectively (Figure 7a–c).

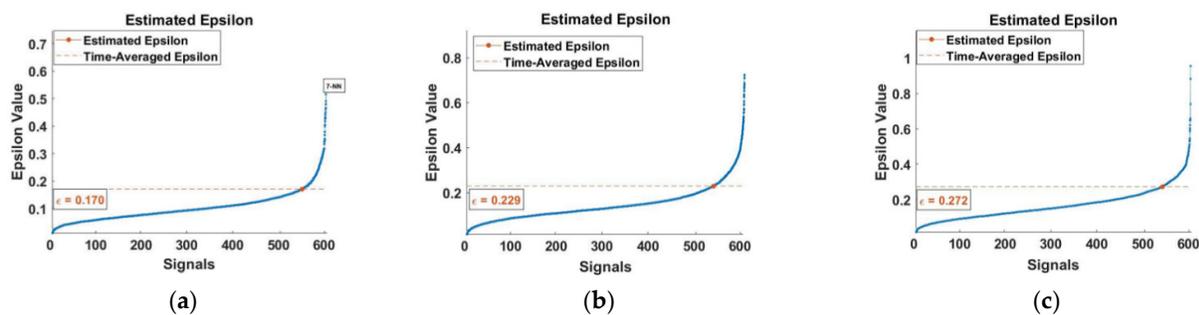


Figure 7. Estimated epsilon, (a) for bearing number 1, (b) for bearing number 2, and (c) for bearing number 3.

The DBSCAN method to classify bearing fault was constructed and tested to predict the conditions of the bearings, then the performance of the classification was evaluated. The results are shown in Figure 8 for the three bearing numbers. Figure 8a, represents two classes: the first one being the healthy state with 568 signals, and the second one illustrating the first stage of the defect, starting from signal number 569. Figure 8b, displays three classes, the first class being the healthy one with 560 signals, the second one representing the first stage of the defect, starting with the signal number 561, and third class continuing from signals number 577 to the end and representing a severe defect. Figure 8c illustrates three classes, the first one showing the healthy state with 480 signals, the second class goes

from the first stage of the defect, starting with the signal number 481, and the third class starting from signals number 516 to the end represents a severe defect.

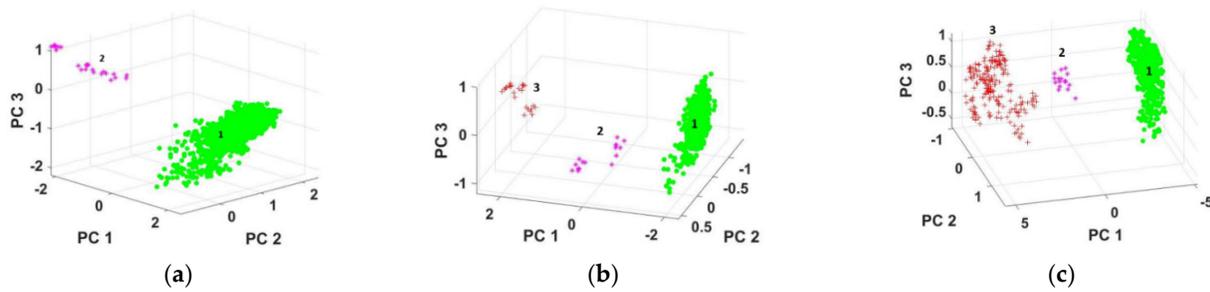


Figure 8. DBSCAN classification, (a) for case number 1, (b) for case 2, and (c) for case number 3.

The results were proportionally correlated to the amplitude, size of defect, value of epsilon, and the detected classes. The size of defect increased with the rise of the Eps value resulting in an augmentation of the classes number. The performance of the proposed method for calculating the bearing health indicator is based on the fusion of the three features, standard deviation, RMS, and peak, which are popular choices for bearing health indicators, then extracting from them the first two principal components.

In Figure 9, the predicted HI and true HI of rolling bearing are demonstrated. As a result, LSTM regressions for the predicted HI and true HI for testing cases 1–3 are shown in Figure 9a–c, respectively. In the third case, the LSTM regression of the predicted HI and true HI were close to each other which indicated that this case was better than cases 1 and 2.

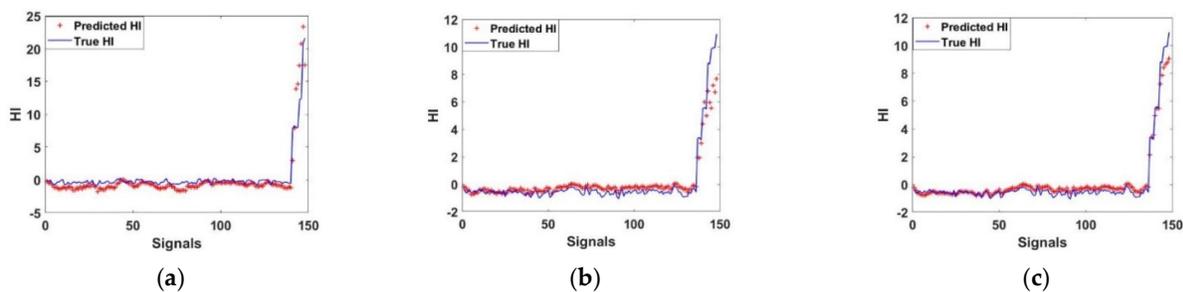


Figure 9. LSTM Regression, (a) case 1, (b) case 2, (c) case 3.

Figure 10 shows the results of the predicted RUL and the true RUL value, where the error between the predicted and true was minimal. To achieve RUL prediction, suitable reliability indicators must be established. Based on the current mapping feature set, the reliability assessment model was firstly utilized to evaluate the precision of the mapping time domain. Then, to build the mathematical exemplification relationship among the component degradation procedure parameters and the prediction RUL model, an LSTM algorithm was utilized in this section. It was established that the accuracy indicator attenuation path was determined prior to the current life and modified to be a fraction of prediction after the present life [47].

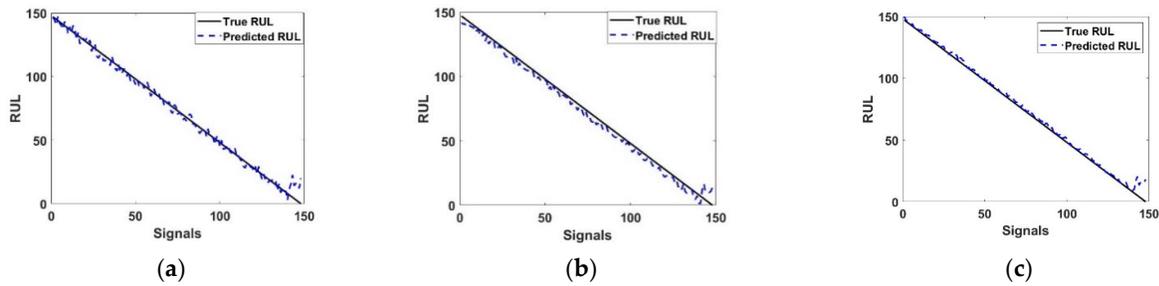


Figure 10. RUL prediction of LSTM in (a) case 1, (b) case 2, (c) case 3.

Case 3 represented in Figure 10c confirms the results shown in Figure 9c, meaning that predicted and true RUL were linear. Moreover, this case represented bearing numbers 2 and 3 which are the best to predict RUL because they have the same degradation state with the three classes of defect.

6. Discussion

To avoid the volatility caused by the trend prediction of the curve by the LSTM method, the value was taken as the evaluation standard, as shown in Table 3. The next phase comprised validation of the trained and tested models. The NARX model’s results shown in Figure 11a,b, present respectively the final results of the model obtained from case 3. Then, the comparison of the performance was done between the LSTM model and the NARX one. The mean absolute error (MAE), the root mean square error (RMSE), and accuracy were adopted to evaluate the performance of the two different prediction models.

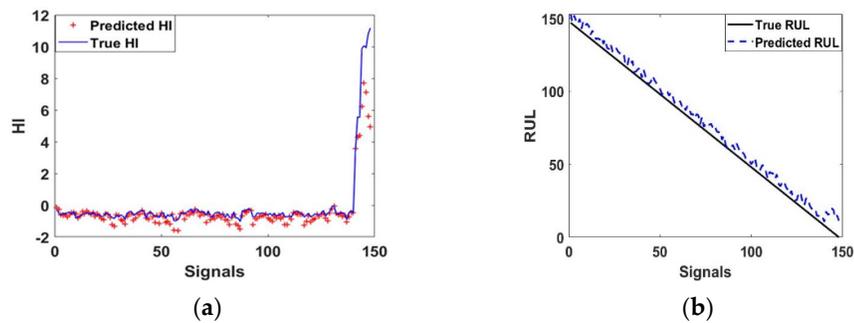


Figure 11. (a) HI Regression, (b) RUL prediction.

As shown in Table 3, the LSTM method case 3 with the proposed HI has the minimum MAE value compared with NARX, and the average MAE value for the LSTM method on datasets of different percentages was 0.007. The NARX method has the biggest MAE with an average value of 0.03. Compared with MAE, the RMSE of the prediction models was larger. For the LSTM method, the average RMSE on different testing datasets was 0.04. Moreover, this metric for the NARX method even reached 0.4. The accuracy value for the LSTM method on datasets of different percentages was 93%. The NARX method has the smallest accuracy with the kurtosis spectral like a health indicator with an average value of 86%. These values showed that the effectiveness of the LSTM method remaining useful life prediction is promising. The elapsed time is the time required by the algorithm to give a response. The LSTM method gave a faster response compared with NARX method.

Table 3. Evaluation of predicted useful life.

	Data	Methods	LSTM	NARX
Proposed HI	Experimental Case 1	RMSE	0.05	0.3
		MAE	0.01	0.02
		Accuracy %	89	87
		Time Elapsed	49 s	52 s
	Experimental Case 2	RMSE	0.3	0.6
		MAE	0.01	0.02
		Accuracy %	85	81
		Time Elapsed	49 s	52 s
	Experimental Case 3	RMSE	0.04	0.4
		MAE	0.007	0.03
		Accuracy %	93	90
		Time Elapsed	49 s	52 s
Kurtosis Spectral Case 3	RMSE	0.1	0.4	
	MAE	0.02	0.03	
	Accuracy %	86	85	
	Time Elapsed	49 s	52 s	

7. Conclusions

This paper proposed the LSTM method to extract high-quality degradation models and predict remaining useful life of rolling bearings from their vibration signals. The extracted bearing vibration signal features included nine time-domain features that were used in the detection chain. The health indicator was constructed from the fusion of three-time domain features: RMS, Std and Peak.

The results showed that the suggested method can adjust to varying operating conditions. The LSTM method was adopted for bearing degradation state and prediction of the remaining useful life. The performance of the proposed method on the dataset was compared with the NARX method whom accuracy is 90%. The results present the superiority and effectiveness of the proposed methodology from the RMSE, MAE and their accuracy are 0.04, 0.007 and 93%, respectively.

The objective of bearing fault diagnosis is to consolidate an efficient real-time condition monitoring and recognition system to make continuous production over predictive maintenance. In order to realize this goal, there are two main topics for future work:

- (1) The fault diagnosis model requires the performance of continual and rapid diagnosis situated on the vibration signals extracted in real-time.
- (2) The good performance was maintained under various working conditions, and the application of the fault diagnosis model must be generalized under different load and noise conditions.

As future work, we plan to examine the use of other algorithms for resolving bearing remaining useful life prediction issues and set up cooperative predictions under different working conditions.

Author Contributions: All authors contributed to the main idea of this paper. Conceptualization, H.H. and X.C.; methodology, H.H.; writing—original draft preparation, H.H.; writing—review and editing H.H., L.R. and X.C.; supervision, X.C., L.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Huang, G.; Li, H.; Ou, J.; Zhang, Y.; Zhang, M. A reliable prognosis approach for degradation evaluation of rolling bearing using MCLSTM. *Sensors* **2020**, *20*, 1864, doi:10.3390/s20071864.
2. Short, M.; Twiddle, J. An Industrial Digitalization Platform for Condition Monitoring and Predictive Maintenance of Pumping Equipment. *Sensors* **2019**, *19*, 3781, doi:10.3390/s19173781.
3. Lu, C.; Wang, Z.; Zhou, B. Intelligent fault diagnosis of rolling bearing using hierarchical convolutional network based health state classification. *Adv. Eng. Inform.* **2017**, *32*, doi:10.1016/j.aei.2017.02.005.
4. Li, H.; Wang, W.; Huang, P.; Li, Q. Fault diagnosis of rolling bearing using symmetrized dot pattern and density-based clustering. *Meas. J. Int. Meas. Confed.* **2020**, *152*, 107293, doi:10.1016/j.measurement.2019.107293.
5. Xu, G.; Hou, D.; Qi, H.; Bo, L. High-speed train wheel set bearing fault diagnosis and prognostics: A new prognostic model based on extendable useful life. *Mech. Syst. Signal Process.* **2021**, *146*, 107050, doi:10.1016/j.ymssp.2020.107050.
6. Kong, X.; Yang, J. Remaining Useful Life Prediction of Rolling Bearings Based on RMS-MAVE and Dynamic Exponential Regression Model. *IEEE Access* **2019**, *7*, 169705–169714, doi:10.1109/ACCESS.2019.2954915.
7. Saidi, L.; Ali, J.B.; Bechhoefer, E.; Benbouzid, M. Wind turbine high-speed shaft bearings health prognosis through a spectral Kurtosis-derived indices and SVR. *Appl. Acoust.* **2017**, *120*, 1–8, doi:10.1016/j.apacoust.2017.01.005.
8. Brusa, E.; Bruzzone, F.; Delprete, C.; di Maggio, L.G.; Rosso, C. Health indicators construction for damage level assessment in bearing diagnostics: A proposal of an energetic approach based on envelope analysis. *Appl. Sci.* **2020**, *10*, 8131, doi:10.3390/app10228131.
9. Yang, Y.; Dong, X.J.; Peng, Z.K.; Zhang, W.M.; Meng, G. Vibration signal analysis using parameterized time-frequency method for features extraction of varying-speed rotary machinery. *J. Sound Vib.* **2015**, *335*, 350–366, doi:10.1016/j.jsv.2014.09.025.
10. Wang, D.; Tsui, K.L.; Miao, Q. Prognostics and Health Management: A Review of Vibration Based Bearing and Gear Health Indicators. *IEEE Access* **2017**, *6*, 665–676, doi:10.1109/ACCESS.2017.2774261.
11. Li, X.; Jiang, H.; Xiong, X.; Shao, H. Rolling bearing health prognosis using a modified health index based hierarchical gated recurrent unit network. *Mech. Mach. Theory* **2019**, *133*, 229–249, doi:10.1016/j.mechmachtheory.2018.11.005.
12. Li, S.; Chen, Z.; Liu, Q.; Shi, W.; Li, K. Modeling and Analysis of Performance Degradation Data for Reliability Assessment: A Review. *IEEE Access* **2020**, *8*, 74648–74678, doi:10.1109/ACCESS.2020.2987332.
13. Okoh, C.; Roy, R.; Mehnen, J.; Redding, L. Overview of Remaining Useful Life Prediction Techniques in Through-life Engineering Services. *Procedia CIRP* **2014**, *16*, 158–163, doi:10.1016/j.procir.2014.02.006.
14. Peng, Y.; Cheng, J.; Liu, Y.; Li, X.; Peng, Z. An adaptive data-driven method for accurate prediction of remaining useful life of rolling bearings. *Front. Mech. Eng.* **2018**, *13*, 301–310, doi:10.1007/s11465-017-0449-7.
15. Mosallam, A.; Medjaher, K.; Zerhouni, N. Data-driven prognostic method based on Bayesian approaches for direct remaining useful life prediction. *J. Intell. Manuf.* **2016**, *27*, 1037–1048, doi:10.1007/s10845-014-0933-4.
16. Lu, C.; Wang, S. Performance Degradation Prediction Based on a Gaussian Mixture Model and Optimized Support Vector Regression for an Aviation Piston Pump. *Sensors* **2020**, *20*, 3854, doi:10.3390/s20143854.
17. Satishkumar, R.; Sugumaran, V. Estimation of remaining useful life of bearings based on Support Vector Regression. *Indian J. Sci. Technol.* **2016**, *9*, 339–349, doi:10.17485/ijst/2016/v9i10/88997.
18. Koopialipoor, M.; Fahimifar, A.; Ghaleini, E.N.; Momenzadeh, M.; Armaghani, D.J. Development of a new hybrid ANN for solving a geotechnical problem related to tunnel boring machine performance. *Eng. Comput.* **2020**, *36*, 345–357, doi:10.1007/s00366-019-00701-8.
19. Hu, Y.W.; Zhang, H.C.; Liu, S.J.; Lu, H.T. Sequential Monte Carlo Method Toward Online RUL Assessment with Applications. *Chin. J. Mech. Eng.* **2018**, *31*, 1–12, doi:10.1186/s10033-018-0205-x.
20. Wang, Y.; Zhao, Y.; Addepalli, S. Remaining useful life prediction using deep learning approaches: A review. *Procedia Manuf.* **2020**, *49*, 81–88, doi:10.1016/j.promfg.2020.06.015.
21. Cali, Ü.; Lange, B.; Dobschinski, J.; Kurt, M.; Ernst, B. Artificial neural network based wind power forecasting using a multischeme ensemble prediction model. *Eur. Wind Energy Conf. Exhib.* **2008**, *1*, 423–432.
22. Prappacher, N.; Bullmann, M.; Bohn, G.; Deinzer, F.; Linke, A. Defect Detection on Rolling Element Surface Scans Using Neural Image Segmentation. *Appl. Sci.* **2020**, *10*, 3290, doi:10.3390/app10093290.
23. Zhang, A.; Wang, H.; Li, S.; Cui, Y.; Liu, Z.; Guanci, Y.; Hu, J. Transfer Learning with Deep Recurrent Neural Networks for Remaining Useful Life Estimation. *Appl. Sci.* **2018**, *8*, 2416, doi:10.3390/app8122416.
24. Tian, Q.; Wang, H. An ensemble learning and RUL prediction method based on bearings degradation indicator construction. *Appl. Sci.* **2020**, *10*, 346, doi:10.3390/app10010346.
25. Leontaritis, I.J.; Billings, S.A. Input-output parametric models for non-linear systems Part I: Deterministic non-linear systems. *Int. J. Control* **1985**, *41*, 303–328, doi:10.1080/0020718508961129.
26. Ruano, A.E.; Fleming, P.J.; Teixeira, C.; Rodríguez-Vázquez, K.; Fonseca, C.M. Nonlinear identification of aircraft gas-turbine dynamics. *Neurocomputing* **2003**, *55*, 551–579, doi:10.1016/S0925-2312(03)00393-X.
27. Ma, Y.; Liu, H.; Zhu, Y.; Wang, F.; Luo, Z. The NARX Model-Based System Identification on Nonlinear, Rotor-Bearing Systems. *Appl. Sci.* **2017**, *7*, 911, doi:10.3390/app7090911.
28. Liu, X.; Lu, C.; Liang, S.; Godbole, A.; Chen, Y. Vibration-induced aerodynamic loads on large horizontal axis wind turbine blades. *Appl. Energy* **2017**, *185*, 1109–1119, doi:10.1016/j.apenergy.2015.11.080.

29. Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In Proceedings of the Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, Portland, OR, USA, 2–4 August 1996; pp. 226–231, doi:10.1.1.121.9220. Available online: <https://citeseerx.ist.psu.edu/viewdoc/summary?> (accessed on 9 April 2021).
30. Zhang, J.; Zeng, Y.; Starly, B. Recurrent neural networks with long term temporal dependencies in machine tool wear diagnosis and prognosis. *SN Appl. Sci.* **2021**, *3*, 442, doi:10.1007/s42452-021-04427-5.
31. Li, X.; Parizeau, M.; Plamondon, R. Training hidden markov models with multiple observations—Combinatorial method. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 371–377, doi:10.1109/34.845379.
32. Smagulova, K.; James, A.P. A survey on LSTM memristive neural network architectures and applications. *Eur. Phys. J. Spec. Topics*, **2019**, *228*, 2313–2324, doi:10.1140/epjst/e2019-900046-x.
33. van Houdt, G.; Mosquera, C.; Nápoles, G. A review on the long short-term memory model. *Artif. Intell. Rev.* **2020**, *53*, 5929–5955, doi:10.1007/s10462-020-09838-1.
34. Baek, Y.; Kim, H.Y. ModAugNet: A new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module. *Expert Syst. Appl.* **2018**, *113*, 457–480, doi:10.1016/j.eswa.2018.07.019.
35. Zou, F.; Zhang, H.; Sang, S.; Li, X.; He, W.; Liu, X. Bearing fault diagnosis based on combined multi-scale weighted entropy morphological filtering and bi-LSTM. *Appl. Intell.* **2021**, doi:10.1007/s10489-021-02229-1.
36. Hotait, H.; Chiementin, X.; Rasolofondraibe, L. AOC-OPTICS: Automatic Online Classification for Condition Monitoring of Rolling Bearing. *Processes* **2020**, *8*, 606, doi:10.3390/pr8050606.
37. Kim, J.H.; Choi, J.H.; Yoo, K.H.; Nasridinov, A. AA-DBSCAN: An approximate adaptive DBSCAN for finding clusters with varying densities. *J. Supercomput.* **2019**, *75*, 142–169, doi:10.1007/s11227-018-2380-z.
38. Hotait, H.; Chiementin, X.; Mouchaweh, M.S.; Rasolofondraibe, L. Monitoring of Ball Bearing Based on Improved Real-Time OPTICS Clustering. *J. Signal Process. Syst.* **2021**, *93*, 221–237, doi:10.1007/s11265-020-01571-w.
39. Tian, J.; Morillo, C.; Azarian, M.H.; Pecht, M. Motor Bearing Fault Detection Using Spectral Kurtosis-Based Feature Extraction Coupled with K-Nearest Neighbor Distance Analysis. *IEEE Trans. Ind. Electron.* **2016**, *63*, 1793–1803, doi:10.1109/TIE.2015.2509913.
40. Samanta, B.; Al-Balushi, K.R. Artificial neural network based fault diagnostics of rolling element bearings using time-domain features. *Mech. Syst. Signal Process.* **2003**, *17*, 317–328, doi:10.1006/mssp.2001.1462.
41. Rai, V.K.; Mohanty, A.R. Bearing fault diagnosis using FFT of intrinsic mode functions in Hilbert-Huang transform. *Mech. Syst. Signal Process.* **2007**, *21*, 2607–2615, doi:10.1016/j.ymsp.2006.12.004.
42. Zheng, K.; Yang, D.; Zhang, B.; Xiong, J.; Luo, J.; Dong, Y. A group sparse representation method in frequency domain with adaptive parameters optimization of detecting incipient rolling bearing fault. *J. Sound Vib.* **2019**, *462*, 114931, doi:10.1016/j.jsv.2019.114931.
43. Ma, J.; Chen, G.; Li, C.; Zhan, L.; Zhang, G.-Z. Rolling Bearing Feature Extraction Method Based on Improved Intrinsic Time-Scale Decomposition and Mathematical Morphological Analysis. *Appl. Sci.* **2021**, *11*, 2719, doi:10.3390/app11062719.
44. Deng, F.; Yang, S.; Liu, Y.; Liao, Y.; Ren, B. Fault Diagnosis of Rolling Bearing Using the Hermitian Wavelet Analysis, KPCA and SVM. In Proceedings of the 2017 International Conference on Sensing, Diagnostics, Prognostics, and Control, SDPC 2017, Shanghai, China, 16–18 August 2017; pp. 632–637, doi:10.1109/SDPC.2017.124.
45. Hotait, H.; Chiementin, X.; Rasolofondraibe, L. Stratégie de Surveillance des Machines Tournantes par Classification en Temps Réel des Signatures Vibratoires. *Congrès Français Mécanique* 26–30 August 2019. Available online: <https://cfm2019.sciences-conf.org/244669> (accessed on 1 October 2019).
46. Le, T.-T. Surrogate Neural Network Model for Prediction of Load-Bearing Capacity of CFSS Members Considering Loading Eccentricity. *Appl. Sci.* **2020**, *10*, 3452, doi:10.3390/app10103452.
47. Qin, Y.; Chen, D.; Xiang, S.; Zhu, C. Gated dual attention unit neural networks for remaining useful life prediction of rolling bearings. *IEEE Trans. Ind. Inform.* **2020**, *1*, doi:10.1109/tii.2020.2999442.