Journal of Engineering Science and Technology
Vol. 16, No. 5 (2021) 3737 – 3756

ESTIMATION OF REMAINING USEFUL LIFE OF BEARINGS USING REDUCED AFFINITY PROPAGATED CLUSTERING

S. SHARANYA[1,*], REVATHI VENKATARAMAN[1], G. MURALI[2]

[1]Department of Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu-603203, India
[2]Department of Mechatronics, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu-603203, India
*Corresponding Author: nice_sharanya@yahoo.co.in

Abstract

Preventive maintenance through predicting the Remaining Useful Life (RUL) of engineering components is considered as a breakthrough in Industry4.0. Researchers are developing prognostic models to predict the machine failures by deploying machine learning and deep learning algorithms. In the era of big data, featuring engineering is a challenging task in building predictive analytic models. The proposed methodology engrosses raw vibration signals of bearings and forms exemplar data through novel Reduced Affinity Propagated (RAP) clustering algorithm. Echo State Network with dynamic connections predicts the RUL from the natural grouping of the data. With these novelties, the developed approach outperforms the other state of art techniques. The model is robust and generic so that it can become a potential candidate for prognosis of higher engineering components.

Keywords: Echo state networks (ESN), Reduced affinity propagated (RAP) clustering, Remaining useful life (RUL), Reservoir unit, Time window.

1. Introduction

The dawn of Industry 4.0 has increased the deployment of advanced engineering systems in time critical and life- saving applications. This very fact seeks the attention of researchers to develop machine health prognostic procedures and methodologies using state of techniques. The streaming real time data, advancements in processing capabilities and new signal processing techniques has opened new ventures in Prognostics and Health Management (PHM) of machineries.

The machine health monitoring is done in three phases namely detection of fault, its diagnosis and prediction of Remaining Useful Life (RUL) [1]. The prediction of RUL, which is one of the prominent prognostic techniques, is very crucial in machine health prognostics. It plays a key role in decisive processes involved in mission planning, fault mitigation and failure analysis [2]. The ISO 13381-1 defines prognostics as the estimated time to failure and risk of one or more existing assets and the prediction of future failure modes [3]. Prognostics is seen as an array of activities to build systems with safety, availability and low-cost maintenance [4].

RUL is the time span during which the equipment is expected to ensure its fullest performance in its intended purpose. The performance of the equipment is tested from the monitored critical variables. Any deviation from its standards will be viewed as an abnormal activity. The prognostic RUL prediction is generally done using model based and analytical based methodologies [5]. The model-based methods are designed using statistical knowledge and computational intelligence on run-to-failure data whereas the analytical approaches (Physics-of-failure) deploys the knowledge on physical property of the material. The former methods gain more popularity, since they do not require domain expertise, but the historic data can precisely model the nature of degradation. The data emanating from the systems may be classified into two types of namely event data and Condition Monitored (CM) data.

The major challenge in manipulating the event data is the non- availability of failure data since critical assets cannot be run-to-failure. So, the natural solution would be to use CM data extracted from signals or operational profiles of equipment. The CM data is further bifurcated into direct CM and indirect CM. The direct CM describes the system by direct data to reach the predefined threshold level while the latter gives a partial indication of the system state through its data. Based on the type of data acquisition (direct or indirect), the system models are also classified as direct model and indirect model [6]. Prediction of RUL using indirect data is done by integrating CM data and model that describes the data. The hybrid models combine the merits of physical models by using degradation status and the historical data.

The model-based methods gained more significance after the advent of machine learning and artificial intelligence. The prognostics have taken a giant leap after the deployment of extreme learning and deep learning techniques since they eliminated the tedious process of feature engineering. These methods are much valued as they can be deployed over both raw and processed signal data. In addition to this, the sensor data can effectively disclose the underlying correlation, thereby aiding the computation of RUL [7]. It is no wonder that the robustness of machine learning algorithms has found them momentous locus in almost every field today.

This work aims to develop a data driven model that estimates the RUL of rotating machinery namely bearing with run to failure data using vibration signals. Bearings are indispensable part of heavy mechanical and civil structures, as they reduce friction between moving parts. Though the bearings are designed for longevity, improper practices and standard wear and tear invite premature bearing failures. The consequences bearing failure may affect the adjacent components like shaft and housings. On the bottom line, frequent bearing failures will heavily cost the industries in terms of reduced operational time and increased downtime; declined performance; amplified maintenance costs and may weaken operator safety. Detecting bearing failure at very initial stage of deterioration by capturing its health trend is a more appropriate choice for industries [8]. It is evident that degradation of bearing is not a sudden phenomenon.

The CM data of the normal bearing may show some random deviations in the initial degraded stage and eventually results in irreversible deterioration of the bearing performance. This trend is well portrayed by vibration signals that can be easily measured through accelerometers mounted in appropriate positions. The simplicity of mounting hardware and rich statistical features makes the vibration signals a potential candidate for bearing health prognostic [9, 10].

This article considers the data emanated from vibration signals to design a novel approach, using Reduced affinity propagated clustering and Recurrent Neural Network (RNN) to predict the RUL of bearings. The major research contributions of this work include:

- Formation of RAP clusters from the temporal vibrational data.
- Choosing time scaled exemplar points based on context-based likelihood.
- Hybridizing the Self Organising Maps (SOM) to form dynamic inhibiting connections from the exemplar points to predict the RUL.

This work will be of immense help in building reliable smart factories where human intervention must be kept at bay. As Industry 4.0 aims to install smartness in all possible areas, this work would assist the maintenance and support team to schedule proactive maintenance activities.

The proposed framework is considered unique of its kind for two major reasons. Firstly, the learning does not happen based on the pre assumed features. Secondly, the exemplar points will act as a representative candidate for signals in particular time frame. The organisation of the paper is: Section 1 introduces the problem domain and brief literature is presented in Section 2. The proposed novel RAP clustering algorithm and estimation of RUL is discussed in Section 3 and 4 respectively. Section 5 describes the formation of dynamic connections and experimental procedure is detailed in Section 6. Section 7 concludes the work with its future scope.

2. Prognostics: A Brief Review

The prognostic methodologies are classified into i) Model based approaches ii) Data driven approaches and iii) hybrid approaches. Literature in prognostics of RUL is inclined towards building systems using hybrid methodologies, which indicates the importance of statistical and domain expertise in the field.

Construction of health indicators in machineries should reflect the monotonic trends and patterns occur in their degradation phase [11].Various knowledge extraction strategies were developed such as empirical mode decomposition [12], recurrence quantification analysis [13], and adaptive time window approach [14], which uses raw signal data to construct health indicators.

The two main issues in these strategies are that the original data is sensitive to faults only in certain specific conditions: after the degradation is progressed to diagnosable extent and existence of correlation among the features is not exploited. These factors mitigate the chances of early fault predictions. Fusing multiple but correlated features to form a more powerful feature space is attempted by Qiu et al. [15] by using Self Organising Maps, Huang et al. [16] by using dynamic principal component analysis and Lei et al. [17] by using particle filtering. Sparse representation score derived from orthogonal matching is also used as an alternate representation for classical features like RMS, kurtosis etc [18]. However, all the features do not positively contribute to the precise prediction because of non-monotonicity exhibited during the degradation process of the components.

Intelligent techniques like Artificial Neural Networks [19], Extreme learning Machines [20], deep Convolutional Neural Network (CNNs) [21], Support Vector Machines [22], Recurrent Neural Networks [23], Adaptive Hidden Markov Models [24], bidirectional Long Short-Term Memory (LSTM) [25] were used in prognostics of RUL. A detailed study on literature of prognostics reveals that ANNs were predominantly used in estimation of RULs, because of their time series forecasting nature. Shallow neural networks integrated with Monte Carlo method was deployed on non-linear time series data for forecasting [26]. ANNs were combined with Weibull failure rate function in predicting the life percentage based on vibration data [27].

Multi valued neurons were used to extract highly dynamic patterns for multi-step ahead predictions [28]. Gugulothu et al. [29] used RNNs over multivariate time series data to form normal and degraded patterns in the form of embedding. Integrating homogeneous and heterogeneous machine learning algorithms has given rise to ensemble models in estimation of RUL. Ensemble of neural networks with Kalman filter was proposed by Peel [30]. Integration of Regression Vector Machines, Support Vector Machines, Least Square exponential Fitting, Bayesian linear regression and RNNs were designed by Li et al [31]. Ensembling of optimised Echo state networks with mean variance estimation was used to deal the uncertainties in estimation of RUL [32]. An ensemble of deep belief network was constructed by Zhang et al. [33] that deploys evolutionary multi-objective function to reduce errors.

Not many works were done in using clustering-based RUL estimation. Some of the noteworthy works are enumerated here. Hu et al. [34] used Fuzzy C-means clustering along with wavelet Support Vector Regression over a degradation path model. Subtractive maximum entropy-based clustering methodology is adapted to label the data to fault classes is proposed by Chiu [35] and further estimation of RUL is done by integrating it with extreme learning machines.

Deep learning methods are now dominating the field of RUL prediction. A double CNN model where first model is used to isolate the incipient fault point and second to predict the RUL is proposed by Yang et al. [36]. An encoder-decoder based recurrent neural network is used by Chen et al. [37] to derive the health index values without thresholding. The final RUL is predicted using linear regression. The generic CNN does not consider temporal variations of the signals. Hence RCNN model is deployed to learn the time-based deviations using recurrent layers [38]. This methodology is used to learn various degradation stages from the vibration data.

The detailed literature in RUL prediction indicates that deep learning methods overshadow machine learning algorithms because of their inherent processing power and wards off the tiresome task of feature engineering from the algorithm. Augmenting to this, the selection of improper features may sometimes lead to adverse effects on the algorithm. Also, these methods exhibit improved performance at relatively lower time scale. However, the adaptive learning occurs in multiple layers, which apparently demands high end resources with considerably good processing power. The streaming signal data will further escalate the complexity of deployment of these methods.

This paper proposes a methodology that uses novel reduced affinity propagation clustering those forms exemplar points based on availability and responsibility of the degradation-based vibration signals. The exemplar points, representing the colossal signal characteristic are fed into the Echo State Networks (ESN) which predicts the RUL of the component. The proposed RAP clustering algorithm reduces substantial number of computations by reorienting the signals into exemplar points, without loss of generality. This method naturally learns from the time-based signals without suffering from the curse of dimensionality.

3. Clustering by Passing Messages

The classical Affinity Propagated (AP) clustering methodology works on the dogma of insinuating the exemplar points among the other data points by message passing mechanism. The AP clustering relies on the real valued similarities among the nodes or data points.

Consider $D_P=\{1, 2, 3, ..., n\}$, the real valued data points in the vectored feature space $\{Dp\ (i, j)\}N\times N$. The affinity measure $A(i, j)$ between the nodes i and j is established by mutual handshaking of responsibility $r(i, j)$ and availability $av(i, j)$ messages among the nodes $r(i, j)$ is the evidence of j being exemplar point to i and $av(i, j)$ is the evidence sent by the candidate exemplar to every data point i. The process is repeated till a strong and stable exemplar node could be designated as a vintage of the so formed cluster. This is illustrated in Figs. 1(a) and (b). Equation (1) gives the expression for estimating the likelihood of exemplar point [39].

$$\max : av(i, j) + r(i, j) \forall i, j \tag{1}$$

For instance, consider D_{Pi} be a node whose affinity measure is $A(i, D_{Pi})$. The exemplar points for D_{Pi} are obtained through maximizing the objective function F through Eq. (2).

$$Max : F(D_P; A) = e^{\sum_{i=1}^{N} A(i, D_{Pi}) + \sum_{j=1}^{N} \log G(D_P)} \tag{2}$$

The range of $D_p=\{1, 2, 3, \ldots, n\}NxN$ and G_j is termed as coherence constant. The first term of Eq. (2) is the affinity measure of node i belonging to its exemplar point E_i. The performance of AP clustering depends largely on hyper-parameter termed self- preference P_j, which is estimated from Eq. (3).

$$p_j = A(j,j) \in A(i,j)_{NxN}\,(i,j = 1,2,3,...,N) \quad (3)$$

The measure P_j is the likelihood of the node j being chosen as an exemplar point for its corresponding cluster. The preference set P comprises of the self- preference values of all the nodes as shown in Eq. (4).

$$P = \{p_1, p_2, p_3, ..., p_N\}_{NxN} \quad (4)$$

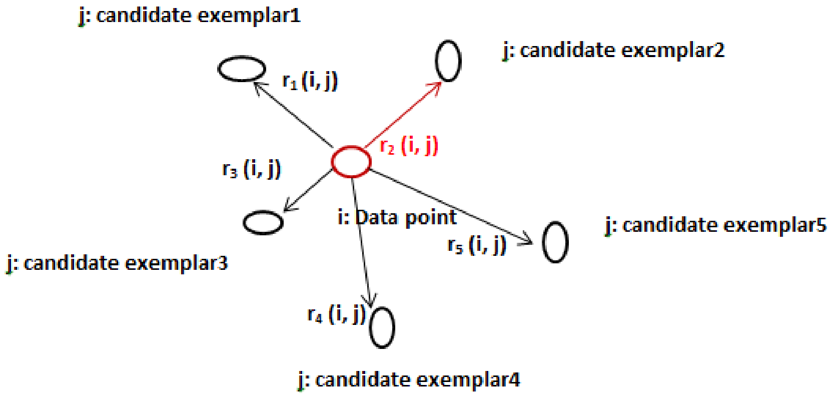

Fig. 1(a). Passing responsibility messages for choosing exemplar.

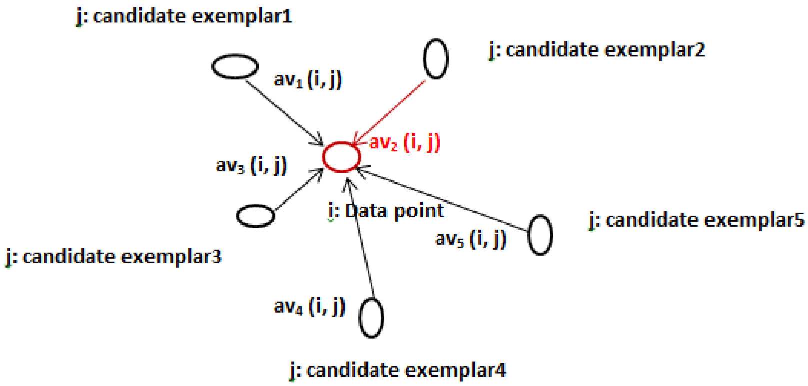

Fig. 1(b). Passing availability messages for choosing exemplar.

The value of P_j is a dominant factor in determining the quality of clusters. A sub optimal low value of P_j will lead to the formation of fewer clusters and high value will lead to more clusters, which will eventually miss the extraction of important correlations between nodes. Confining the P_j value within an upper and lower bound [40], iteratively adjusting the P_j value [41] and incrementing or decrementing P_j values [42] are few notable methods to arrive optimal value. Also, the number of clusters corresponds to the preference points. Hence this method does not confine the number of clusters to a pre-defined value, making the algorithm best suited for online applications. However, deducting a method to fix the preference value is beyond the scope of this work.

3.1. Reduced affinity propagation clustering

The classical AP clustering has two major pitfalls namely choosing preference value for exemplar points and continuous message passing between nodes. The work focuses on estimating the RUL of bearings with low-cost computation but with increased accuracy by aggregating local data. The RAP clustering serves the objective by incorporating:

- Time scaled data: fixing a time window frame.
- Forming temporal exemplar points from degradation signals.

The estimation of responsibility and availability in RAP algorithm is done through estimation of similarities based on the Context Based Multi-layered Bayesian Inference (CBMBI) predictive analytic framework [43]. This predictive analytic technique roots from Mismatch Negativity (MMN), a physiological process that happens inside human brain that estimates the evidence of novel events from the signals transmitted by the sensory organs [44].

In accordance with Newton's free energy principle, the human brain changes its state to supress the free energy generated by the surprise events. This change in state is elicited by the accumulation of sensory signals, which contributes to the free energy. This is the MMN phenomenon. This could be adopted in detecting novel and abnormal events that occurs in condition monitoring of equipment. An upper boundary on the accumulation of evidence is imposed by introducing a novel hyper parameter, Context (C) to mitigate the over fitting effect. The estimation of evidence is done from Eq. (5).

$$g(f_i, S_i; C) = S_i \frac{1}{2\pi} \exp(-I_i / 2)^2 \tag{5}$$

where I_i the input vector of signal data and S_i is the predicted state information of previous state. The I value ranges between *[*1*, C]*.The magnitude of the predicted value of $g(I_i, S_i, C)$ will correspond to the magnitude of the novel event and magnitude is denoted by y of the input cause vector C_V .

$$P(Y; C_V) = P(y \mid C_V) P(C_V) \tag{6}$$

Determining the value of y, provisions the classification of health state of the machine into multiple states based on the significance of the equipment under study. Estimating the $P(y; C_V)$ for every node reduces the flooding of messages to form exemplar points, as the transmission happens only after the sufficient value for y is obtained. This is given in Eq. (6). This aggregated event limits the number of messages thus reducing the complexity. Limiting the messages will increase the efficiency for the RAP in two folds: reduces the complexity and number of transmissions. The exemplar points (E_i) will be formed from the temporal signals and will act as representative of the behaviour of degradation signals.

$$E_i = \{E_0, E_{1,} E_2, ..., E_N\}_{NxN}, i \in t_N \tag{7}$$

The set E_i contains all the exemplar points in the given time domain t_N as mentioned in Eq. (7). Thus, this novel non-linear, non- parametric methodology captivates the signal characteristics through exemplar points. The robustness of the model is enhanced by fixing the self-preference value (p_j) of each exemplar by a fixed size, moving time window. As the estimation of RUL is done using Echo State Network (ESN), it is important to consider the correlation between the data points. Hence, fixing p_j values and forming exemplar points is dependent on the movement of the time window. The time window size (w) is also a significant hyper parameter, whose value is set to the scale that exhibits the smallest operational characteristic profile from the signal data.

4. Estimation of RUL from Temporal Data

RNNs are excellent candidate in estimating RUL, since their computations are powered by time information. The recurrent edges in RNN have 2 types of self-connections: one from previous layer and another from themselves in the previous state (iteration). The description of RNN is given in Fig. 2.

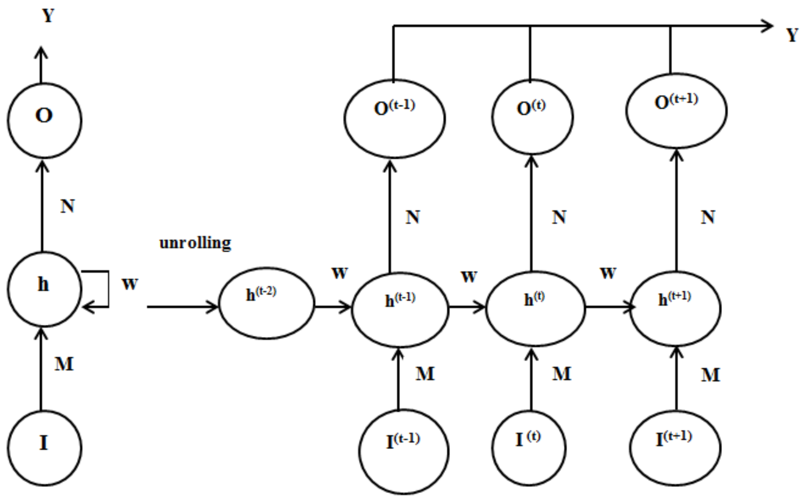

Fig. 2. Single output Recurrent Neural Network (RNN).

The input at time instant t (I (t)) is fed inside the hidden layer $h(^t)$ acted upon by the weight (W), which is homogeneously maintained as same value. M and N are parameterized weights, acting between input-hidden layer and hidden-output layer, respectively. $O(^t)$ is the net output from the RNN at time index t. As the work focus on predicting the RUL, which is a single value rather than a sequence, the designed RNN will cumulatively output a single value Y, which is the normalised probability over the obtained output at various time indices between $[0, \pi]$.

$$i^{(t)} = j + Wh^{(t-1)} + MI^{(t)} \tag{8}$$

$$h^{(t)} = F(MI^{(t)} + Wh^{(t-1)}) \tag{9}$$

$$O^{(t)} = k + Mh^{(t)} \tag{10}$$

$$Y = norm\_O^{(t)} \tag{11}$$

The mathematical descriptions of RNN are given in Eq. (8-11) [45].The main drawbacks of Vanilla RNN are the exploding and vanishing gradient problems. ESN alleviate this by replacing the hidden layer by internal reservoir units with dynamic lateral connections as given in Fig. 3.

The fixed size time window frame moves over the temporal exemplar values $E_i = \{E_0, E_1, E_2, ..., E_N\}$ where $0 \le i \le N$, in overlapping fashion on a scale of $N/2$, so that important patterns are not overlooked. T_{i+1} is the prediction for $(i+1)^{th}$ exemplar points using RNN.

$$T_0 = (E_0, E_1, E_2, ..., E_{N-1}) \tag{12}$$

$$T_1 = (E_{\frac{N-1}{2}}, E_{\frac{N}{2}}, ..., E_{\frac{2N-1}{2}}) \tag{13}$$

$$T_i = (E_{\frac{iw}{2}}, E_{\frac{iw+1}{2}}, ..., E_{\frac{(i+1)w-1}{2}}) \tag{14}$$

$$T_{i+1} = (E_{\frac{(i+1)w}{2}}, E_{\frac{(i+1)w+1}{2}}, ..., E_{(i+2)(w+N)}) \tag{15}$$

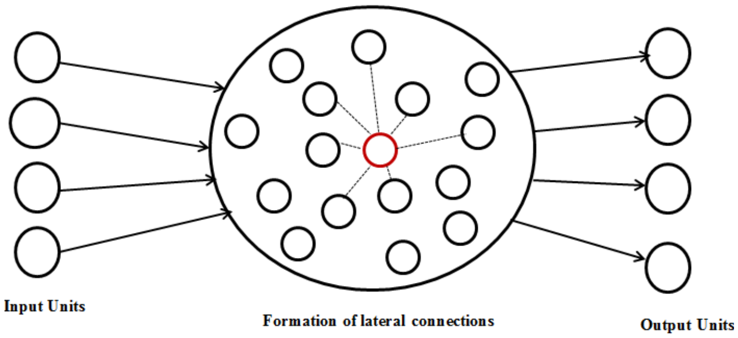

Fig. 3. Dynamic connections in reservoir unit.

Equations (12-15) indicates the time window movement. The predicted RUL from the RNN model will be the normalised value ($norm\_O^{(t)}$) over time period t. The range of $norm\_O^{(t)}$ is $[0, \pi]$ where π is the maximum life duration of the component.

5. Connections through Lateral Inhibitions in Reservoir unit

The classical ESN use fixed connections to activate neurons in the reservoir unit [46]. By exploring the competitive learning scenario as in Self Organising Maps (SOM), the ESN can form lateral connections following Mexican hat function [47]. The primary hyperparameter is the number of neurons in the reservoir unit. The work considers a nominal count of 100 reservoir neurons. In the recurrent SOM, the Best Matching Unit (BMU) that correlate with the input vector $i(n)$ is formed according to the Eq. (16).

$$\|i(n) - hid(n)\| = \min\{\|i - w_h(n)\|\}, \forall w_h \tag{16}$$

where w_h is the weight vector. Reservoir excitation functions are responsible for updating the weights of the neurons in the reservoir unit. This updating is done in correlation with the BMU according to Mexican hat function, which does equal approximations to the matched neurons.

$$r\_ex(n) = exp(\frac{-\|hid_i - BMU_i\|^2}{\vartheta}) \tag{17}$$

The excitation level ($r\_ex(n)$) is estimated using Eq. (17). ϑ is the factor that limits the spread of inhibiting the lateral connections. Weight updating process is in accordance with Eq. (18).

$$w_{n+1} \approx w_n + \alpha r\_ex(n) \tag{18}$$

The term α plays a chief role in convergence of weight in next iteration which in other words resolves vanishing and exploding gradients problem. Thus, the dynamic connections inside the reservoir units are established based on the coherence of the weight vector based on the input. The proposed methodology of RUL prediction in the training phase is given in Fig. 4.

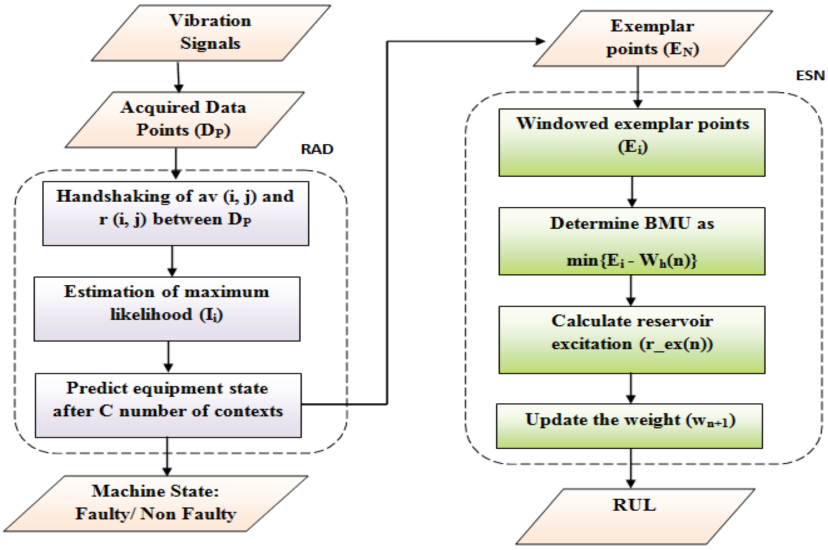

Fig. 4. Remaining Useful Life (RUL) estimation in training phase.

6. Experimental Study

The methodology proposed in this paper is evaluated on a PRONOSTIA, a benchmarked dataset containing vibration and temperature readings from degradation signals of bearings [48] whose experimental set up is shown in Fig. 5. This dataset is more realistic comprising of data samples with inner race, outer race, cage and ball failures of the bearings.

The dataset characterizes the operational profile of the bearings till run-time-to-failure. The vibration signals are measured with 4kN radial force at a rotatory speed of 1800 rpm. 2 types of sensors are deployed over the test bed for capturing vibration and temperature readings.

2 accelerometers (devices that record the vibration signals) are orthogonally placed at outer race of the bearings along vertical and horizontal axes. Type DYTRAN 3035B accelerometer is used to record the vibrations. The temperature sensor (RTD-Resistance Temperature Detector PT100) is installed in a hole in proximity to the outer bearing's ring. The AC motor in the experimental set up runs with a power of 250W with a rotational speed of 2830 rpm. The bearing test bed

has a shaft that drives the inner race of the bearing. The vibration sensors are positioned at right angles through the accelerometers mounted in horizontal and vertical plane. Each sample set of vibration signal has 2560 data points with sampling rate of 10s at duration of 0.1s and the sampling frequency of temperature is 10Hz. The RUL of the bearing is estimated from amplitude of the signals. This dataset contains data collected from ball bearings with outer race diameter of 32mm, inner race diameter of 20 mm and thickness of 7 mm. The bearing is said to operate in failure profile if its vibrational amplitude exceeds 20g (1g = 9.8m / s2). The degradation of the bearings happens naturally without external fault induction. Fig. 6. shows the normal and worn-out bearing. The vibrational pattern depends on the type of bearing used and on the operating condition. Hence there is no uniformity in the degradation nature of the bearings under test.

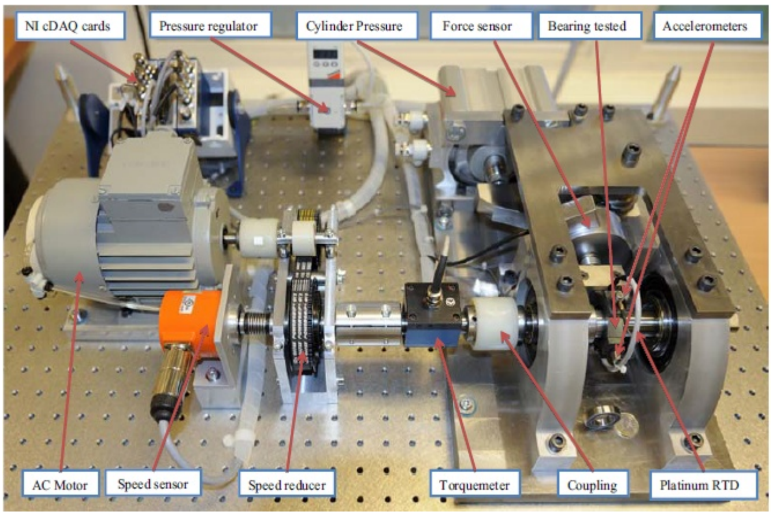

Fig. 5. Experimental system.

Fig. 6. Normal and degraded bearings.

17 bearings are subjected to the following three Operating Conditions (OC) and their signals were recorded as shown in Table 1. The test and training data details are given in Table 2.

Table 1. Operating conditions of bearings.

Operating condition	Load (N)	Speed (rpm)
OC1	4000	1800
OC2	4200	1650
OC3	5000	1500

Table 2. Training and test dataset.

Data sets	OC1	OC2	OC3
Training set	Bearing1_1	Bearing2_1	Bearing3_1
	Bearing1_2	Bearing2_2	Bearing3_2
Test set	Bearing1_3	Bearing2_3	Bearing3_3
	Bearing1_4	Bearing2_4	
	Bearing1_5	Bearing2_5	
	Bearing1_6	Bearing2_6	
	Bearing1_7	Bearing2_7	

The raw vibration signals are providing pool of useful features, which are to be extracted. The raw vibrational signal of bearing1_1 shown in Fig. 7 indicates that the amplitude of the signal increases gradually and has seen its peak value at the end. Studies reveal that frequency range of faulty bearing is {-30g to 30g}, which is relatively very high than that of normal bearing [49]. The vibrational signals can be seen as potential pool of features. Some of the prominent features that could be extracted from the vibrational data is summarised in Fig. 8.

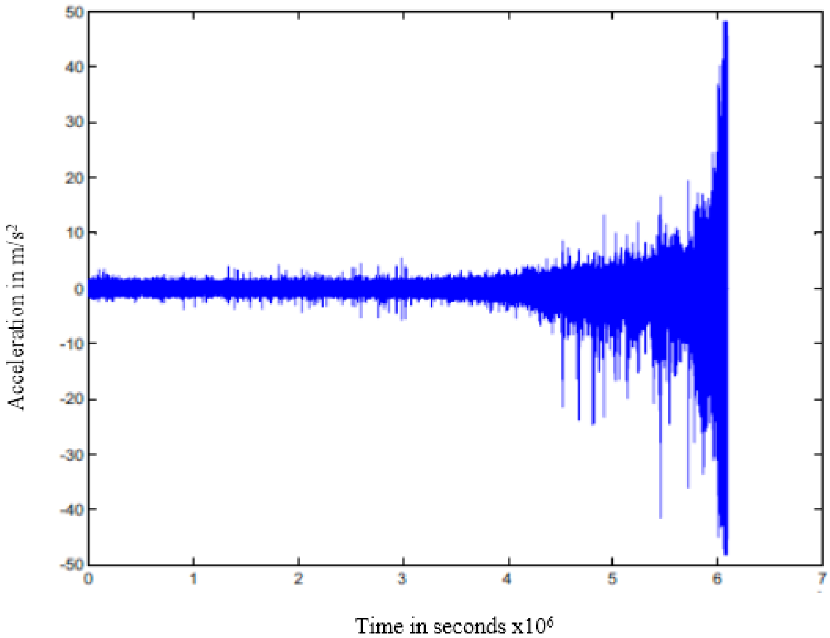

Fig. 7. Raw vibration signal.

(a) Amplitude.

(b) K Factor measured by horizontal accelerometer.

(c) K Factor measured by vertical accelerometer.

(d) Crest factor by horizontal accelerometer ($\times 10^4$).

(e) Crest factor by vertical accelerometer ($\times 10^4$).

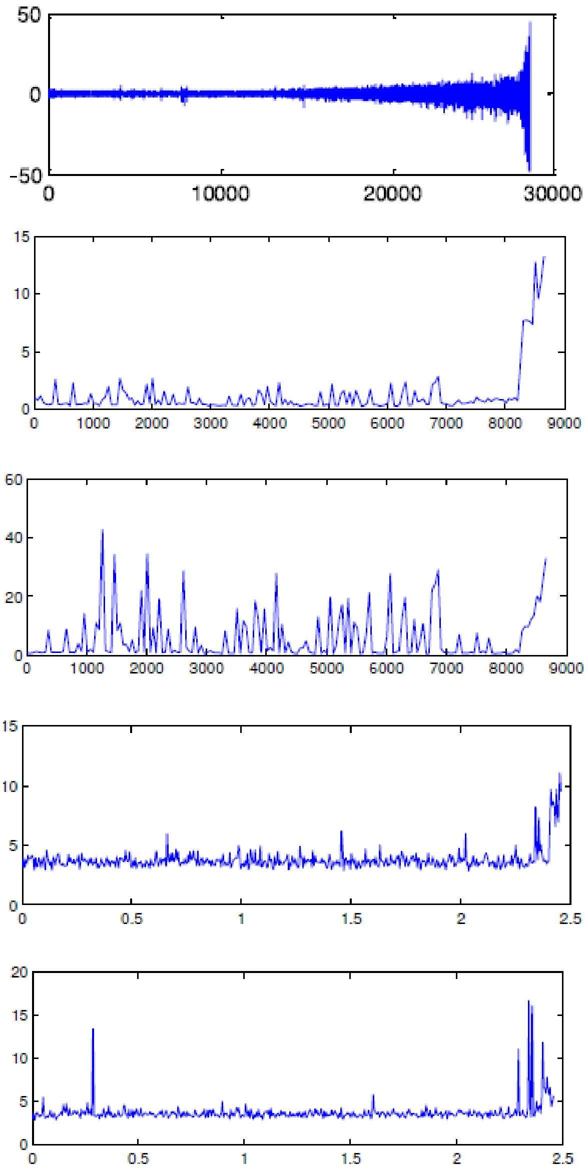

Fig. 8. (a-e) Signal analysis of PRONOSTIA (X axis: Time in seconds × 10^6, Y axis: Acceleration in m/s^2).

It is evident from Fig. 8 that extracting features from exploratory dataset like PRONOSTIA is challenging because of the following signal properties: i) nonlinearity ii) highly unstable iii) low resolution of signal frequency iv) non periodic. The degradation of the bearings is seen as evolution of faults over the lifetime of the bearings. Hence accurate and precise features are pedestal for building of RUL prediction system. In this view, deep learning algorithms fit as natural solutions for domains where manual feature engineering is a challenging task. The proposed RAD algorithm-based RUL prediction relies on deep learning methodology, which has built-in feature engineering that manifests two-fold felicity: it eradicates the effort to learn or extract features and does not overpass the features. The model shown in Fig. 9 is developed with different testing and training phases to reduce computation time.

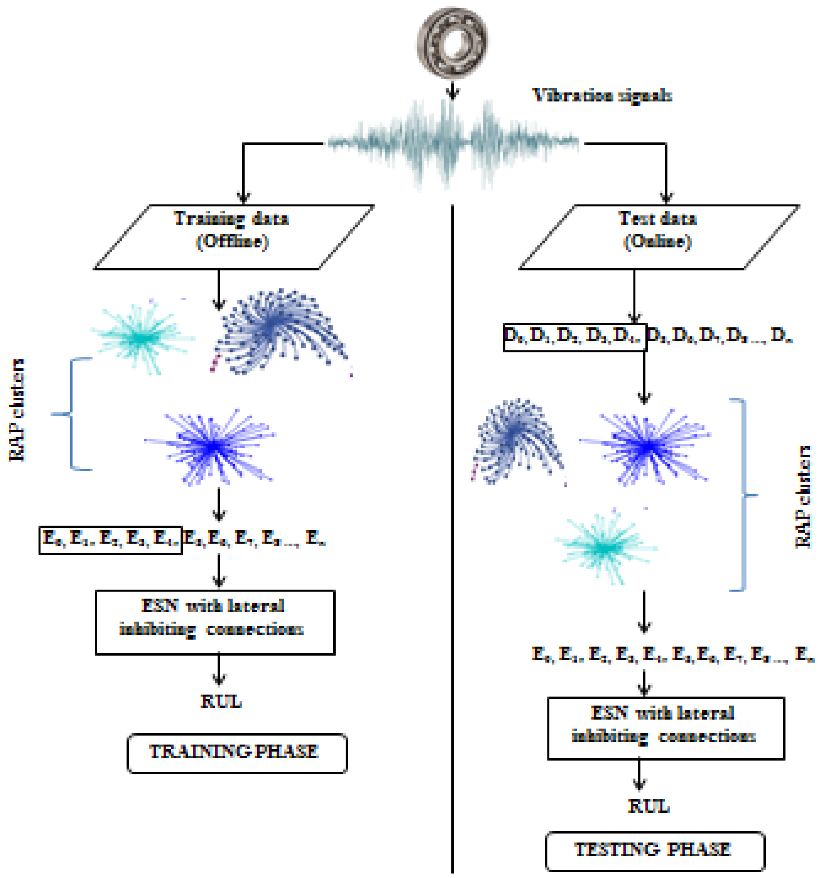

Fig. 9. Operations in training and testing phase in estimation of RUL using RAP.

In the training phase, the ESN network trains the exemplar points formed from RAP clustering in a time scaled fashion. On the other hand, the testing phase draws a distinction by treating raw data points through a time window to form exemplar points. This imposes superior chance for the model to vintage better accuracy at reasonably reduced training. The testing phase witnesses windowing technique prior to the formation of exemplar points to provide adequate evidence. The analysis of the data points was carried out with static window size of 10.

Exploration of the experimental results

The run to failure acceleration data of the bearings has 6 datasets (Bearing1_1, Bearing2_1, Bearing3_1, Bearing1_2, Bearing2_2, Bearing3_2) for training the model. The signal analysis of Bearing1_1 shown in Fig. 10 portrays that the acceleration of signals is more focussed towards negative direction. The dense nature of the plot is a clear indication of the richness of the dataset. The signals of Bearing 1_1 (training) and Bearing 1_3 (test) shows gradual fault progress.

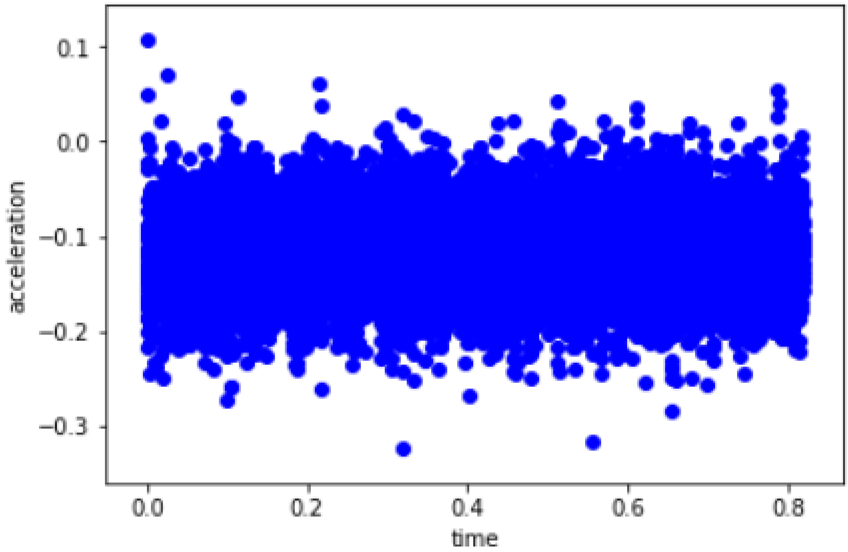

Fig. 10. Data distribution of Bearing1_1.

The test dataset (Bearing1_3,Bearing1_4, Bearing1_5, Bearing1_6,Bearing1_7, Bearing2_3, Bearing2_4,Bearing2_5, Bearing2_6, Bearing2_7, Bearing3_3) comprises of acceleration degradation signals. The vibrational signals are recorded from origin till total faulty condition. The degradation behaviour of the bearings is non-monotonic without any uniformity in their distribution. The analysis of fault development of Bearing 1_5 and Bearing 1_6 exhibits initial swift degradation but tend to be steady in later stages. On the other hand, Bearing1_7 and Bearing 1_8 shows persistent degradation pattern in the run-to-failure lifecycle.

The experimental results of the bearings in training phase do not adhere to any common pattern. Some of them showed a trend in degradation stage (Bearing 1_1) while others were characterized by high feature values leading to sudden failures. So, the validation of the estimated RUL (RUL_E) by the proposed methodology is done by quantifying their accuracy score (A_C), Percent Error ($PerEr_i$) and final score (S_C) in accordance with IEEE PHM 2012 challenge [48] and they are described as Eq. (19, 20, 21).

$$PerEr_i = \frac{RUL_{AT} - RUL_E}{RUL_{AT}} x100 \tag{19}$$

$$A_C = \begin{cases} \exp^{-\ln(0.5)(PerEr_i/5)} & if PerEr_i \leq 0 \\ \exp^{\ln(0.5)(PerEr_i/5)} & if PerEr_i > 0 \end{cases} \tag{20}$$

$$S_C = \frac{\sum_{i=1}^{N} A_{C_i}}{N} \tag{21}$$

Table 3 illustrates the performance of the proposed methodology employing RAP algorithm for natural grouping of vibrational data of the test bearing set as mentioned in Table 2. The negative error values portrayed by Bearing1_5, Bearing 2_4 and Bearing 2_7 are the results of overestimates of RUL_E. Furthermore, the prediction result of the proposed methodology is compared with other state of art techniques [23, 50]. The proposed RAP clustered RUL prediction model has much low error estimates ($PerEr_i$), which gives a direct implication that the model learns better from the raw vibration signals, since there is no overlooking of significant features. Also, the time window scaling of raw vibration signals in training phase and exemplar points in testing phase lowers the computational overhead of ESN. Results of detailed comparisons are tabulated as Table 4. Though the results of the proposed methodology exhibit promising results, it is still competitive.

Table 3. Actual vs. predicted RUL.

Bearing dataset	Actual RUL (s)	Predicted RUL (s)	PerEr$_i$	Ac
Bearing1_3	5730	4561	20.406	0.05
Bearing1_4	2900	2761	4.79	0.51
Bearing1_5	1610	1723	-7.01	2.64
Bearing1_6	1460	1258	13.83	0.14
Bearing1_7	7570	7283	3.79	0.59
Bearing2_3	7530	6853	8.99	0.28
Bearing2_4	1390	1423	-2.37	1.38
Bearing2_5	3090	2642	14.49	0.13
Bearing2_6	1290	1101	14.65	0.13
Bearing2_7	580	651	-12.24	5.45
Bearing3_3	820	761	7.19	0.36
Score (S$_C$)				**1.06**

Table 4. Comparative analysis of proposed methodology with state of art techniques.

Bearing dataset	Percent error (PerEri)				
	Qian et al. [13]	Boccato et al. [47]	Maass et al. [46]	Ali et al. [19]	Proposed method
Bearing1_3	-0.34	-31.76	-1.04	43.28	20.40
Bearing1_4	5.60	62.55	-20.94	62.06	4.79
Bearing1_5	100	-136.03	-278.26	-22.98	-7.01
Bearing1_6	28.08	-32.88	19.18	21.23	13.83
Bearing1_7	-19.55	-11.09	-7.13	17.83	3.79
Bearing2_3	-20.18	44.22	10.49	37.84	8.99
Bearing2_4	8.63	-55.4	51.8	-19.42	-2.37
Bearing2_5	23.30	68.61	28.8	54.36	14.49
Bearing2_6	58.91	-51.94	-20.93	-13.95	14.65
Bearing2_7	5.17	-68.97	44.83	-55.17	-12.24
Bearing3_3	40.24	-21.96	-3.66	3.65	7.19
Average PerEr$_i$	**20.89**	**-21.33**	**-20.31**	**11.70**	**6.04**

7. Conclusion and Future Work

Prediction of RUL is a key factor to schedule preventive maintenance in industries. The proposed methodology analyses raw vibrational signal in its natural perspective to estimate the RUL of bearings. There are 2 major contributions in this work: To begin with, RAP clustered RUL prediction model scales down the vibrational signal by clustering them through message passing. Finally, the lateral dynamic connections inside the reservoir unit are established based on the temporal exemplars formed from RAP clusters. Validation of the model shows its sovereignty over the state of art models. The built model is very generic and robust, thus captivating the industrial researchers to deploy transfer learning to predict the RUL of sophisticated equipment and higher order mechanical systems with data fusion mechanisms. In addition to this, setting the failure threshold and dynamic scaling of time window are subjective matter of interest in prognosis of RUL. This methodology with few domains specific modifications can be used to classify the health state of the equipment under study by constructing health metric from the predicted RUL.

References

1. Lee, J.; Wu, F.; Zhao, W.; Ghaffari, M.; , Liao, L.; and Siegel, D. (2014). Prognostics and health management design for rotary machinery systems-Reviews, methodology and applications. *Mechanical Systems and Signal Processing*, 42(1-2), 314-334.
2. Sankararaman, S. (2015). Significance, interpretation, and quantification of uncertainty in prognostics and remaining useful life prediction. *Mechanical Systems and Signal Processing*, 52. 228-247.
3. Tobon-Mejia, D.A.; Medjaher, K; and Zerhouni, K. (2010). The ISO13381-1 standard's failure prognostics process through an example. *IEEE Prognostics and System Health Management Conference*, University of Macau, Macau, China.
4. Heng, H.; Zhang, S.; Tan, A.C.C.; and Mathew, J. (2009). Rotating machinery prognostics: State of the art, challenges and opportunities. *Mechanical Systems and Signal Processing*, 23(3), 724-739.
5. Okoh, C.; Roy, R.; Mehnen, J.; and Redding, L.(2014). Overview of remaining useful life prediction techniques in through-life engineering services. *Procedia CIRP* 16, 158-163.
6. Valis, D.; Zak, L.; and Pokora, O. (2014). Contribution to system failure occurrence prediction and to system remaining useful life estimation based on oil field data. *Proceedings of the Institution of Mechanical Engineers, Journal of Risk and Reliability*, 1-10.
7. Li, X.; Ding, Q.; and Sun, J.Q. (2018). Remaining useful life estimation in prognostics using deep convolution neural networks. *Reliability Engineering and System Safety*, 172, 1-11.
8. Ahmad, W.; Khan, S.A.; and Kim, J.M. (2018). A hybrid prognostics technique for rolling element bearings using adaptive predictive models. *IEEE Transactions on Industrial Electronics*, 62(2), 1577-1584.
9. Guo, L.; Li, N.; Jia, F.; Yaguo, L.; and Lin, J. (2017). A recurrent neural network based health indicator for remaining useful life prediction of bearings. *Neurocomputing*, 240, 98-109.
10. Yu, J. (2017). Adaptive hidden Markov model-based online learning framework for bearing faulty detection and performance degradation monitoring. *Mechanical Systems and Signal Processing*, 83, 149-162.
11. Camci, F.; Medjaher, K.; Zerhouni, N.; and Nectoux, P. (2013). Feature evaluation for effective bearing prognostics. *Quality and Reliability Engineering International*, 29(4), 477-486.

12. Li, R.; and He, D. (2012). Rotational machine health monitoring and fault detection using EMD-based acoustic emission feature quantification. *IEEE Transaction on Instrumentation and Measurement*, 61(4), 990-1001.
13. Qian Y.; Hu S.; and Yan R. (2013). Bearing performance degradation evaluation using recurrence quantification analysis and auto-regression model. *Proceedings of IEEE International Conference in Instrumentation and Measurement*, Minneapolis, USA, 1713-1716.
14. Xiao, L.; Chen, X.; Zhang, X.; and Liu, M. (2015). A novel approach for bearing remaining useful life estimation under neither failure nor suspension histories condition. *Journal of Intelligent Manufacturing*, 28, 1893-1914.
15. Qiu, H.; Lee J.; Lin, J.; and Yu, G. (2003). Robust performance degradation assessment methods for enhanced rolling element bearing prognostics. *Advanced Engineering Informatics*, 17(3), 127-140.
16. Huang R.; Xi, L.; Li, X.; Liu, R.; Qiu, H.; and Lee, J. (2007). Residual life predictions for ball bearings based on self-organizing map and back propagation neural network methods. *Mechanical Systems and Signal Processing*, 21(1), 193-207.
17. Lei, Y.; Li, N.; Gontarz, S.; Lin, J.; Radkowski, S.; and Dybala. J. (2016). A model-based method for remaining useful life prediction of machinery. *IEEE Transactions on Reliability*, 65(3), 115-123.
18. Nie, Y.; and Wan, J. (2015). Estimation of remaining useful life of bearings using sparse representation method. *Proceedings of* 2015 *Prognostics and System Health Management Conference*, Beijing, China, 1-6.
19. Ali, J.B.; Chebel-Morello, B.; Saidi, L.; Malinowski, S.; and Fnaiech, F. (2015).Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network. *Mechanical Systems and Signal Processing*, 56, 150-172.
20. Javed, K.; Gouriveau, R.; and Zerhouni, N. (2015). A New Multivariate Approach for Prognostics Based on Extreme Learning Machine and Fuzzy Clustering. *IEEE Transactions on Cybernetics*, 45(12), 2626-2639.
21. Xiao, L.; Chen, X.; Zhang, X.; and Liu, M. (2017). A novel approach for bearing remaining useful life estimation under neither failure nor suspension histories condition. *Journal of Intelligent Manufacturing*, 28, 1893-1914.
22. Patil, M.A.; Tagade, P.; Hariharan, K.S.; Kolake, S.M.; Song, T.; Yeo, T.; and Doo, S. (2015). A novel multistage support vector machine based approach for Li ion battery remaining useful life estimation. *Applied Energy*, 159, 285-297.
23. Guo, L.; Li, N.; Jia, F.; Lei, Y.; and Lin, J. (2017). A recurrent neural network based health indicator for remaining useful life prediction of bearings. *Neurocomputing*, 240, 98-109.
24. Yu, J. (2017). Adaptive hidden Markov model-based online learning framework for bearing faulty detection and performance degradation monitoring. *Mechanical Systems and Signal Processing*, 83, 149-162.
25. Elsheikh, A.; Yacout, S. ; and Ouali, M.S. (2018). Bidirectional handshaking LSTM for remaining useful life prediction. *Neurocomputing*, 323, 148-156.
26. Zhou, Q.; Chen, H.; Zhao, H.; Zhang. G.; Yong, G.; and Shen. J. (2016). A local field correlated and Monte Carlo based shallow neural network model

for nonlinear time series prediction. *EAI Endorsed Transactions on Scalable Information Systems*, 3 (8), 1-9.

27. Tian, Z. (2012). An artificial neural network method for remaining useful life prediction of equipment subject to condition monitoring. *Journal of Intelligent Manufacturing*, 23, 227-237.
28. Fink, O.; Zio, E.; and Weidmann, U. (2014). Predicting component reliability and level of degradation with complex-valued neural networks. *Reliability Engineering and System Safety*,121, 198-206.
29. Gugulothu. N.; Vishnu, T.V.; Malhotra, P.; Vig, L.; Agarwal, P.; and Shroff, G. (2017). Predicting remaining useful life using time series embeddings based on recurrent neural networks. *International Journal of Prognostics and Health Management*, 9(1), 1-11.
30. Peel, L. (2008). Data driven prognostics using a Kalman filter ensemble of neural network models. *International Conference on Prognostics and Health Management*, Colorado, USA, 1-6.
31. Li. Z.; Wu, D.; Hu, C.; and Terpenny, J. (2019). An ensemble learning-based prognostic approach with degradation-dependent weights for remaining useful life prediction. *Reliability Engineering and System Safety*, 184, 110-122.
32. Rigamonti, M.; Baraldi, P.; Zio, E.; Roychoudhury, I.; Goebel, K.; and Poll, S.(2017). Ensemble of optimized echo state networks for remaining useful life Prediction. *Neurocomputing*, 281, 121-138.
33. Zhang, C.; Lim, P.; Qin, A.K.; and Tan, K.C. (2016). Multiobjective deep belief networks ensemble for remaining useful life estimation in prognostics. *IEEE Transactions on Neural Networks and Learning Systems*, 28(10), 2306-2318.
34. Hu, Y.T.; Hu, C.H.; Kong, X.Y.; and Zhou, Z.J.(2012). Real-time lifetime prediction method based on wavelet support vector regression and fuzzy c-means clustering. *Acta Automatica Sinica*, 38(3), 331-340.
35. Chiu S.L. (1994). Fuzzy model identification based on cluster estimation. *Journal of intelligent and Fuzzy systems*. 2(3), 267-278.
36. Yang, B.; Liu, R.; and Zio, E. (2019). Remaining useful life prediction based on a double-convolutional neural network architecture. *IEEE Transactions on Industrial Electronics*, 66(12), 9521-9530.
37. Chen, Y.; Peng, G.; Zhu, Z.; and Li, S. (2020). A novel deep learning method based on attention mechanism for bearing remaining useful life prediction. *Applied Soft Computing*, 86.
38. Wang, B.; Lei, Y.; Yan, T.; Li, N.;and Guo, L. (2020). Recurrent convolutional neural network: A new framework for remaining useful life prediction of machinery. *Neurocomputing*, 379, 117-129.
39. Jia, H.; Ding, S.; Meng, L.; and Fan, S. (2014). A density-adaptive affinity propagation clustering algorithm based on spectral dimension reduction. *Neural Computing and Applications*, 25,1557-1567.
40. Givoni; I.E., and Brendan, J. (2009). A binary variable model for affinity propagation. *Neural computation*, 21(6),1589-1600.
41. Li, P.; Ji, H.; Wang, B.; Huang, Z.; and Li, H. (2017). Adjustable preference affinity propagation clustering. *Pattern Recognition Letters*, 85, 72-78.

42. Yang, C.; Bruzzone, L.; Guan, R.; Lu, L.; and. Liang, Y. (2013). Incremental and decremental affinity propagation for semisupervised clustering in multispectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 51(3), 1666-1679.
43. Sharanya, S.; and Venkataraman, R. (2020). An intelligent context based multi-layered Bayesian inferential predictive analytic framework for classifying machine states. *Journal of Ambient Intelligence and Humanized Computing*, 12, 7353-7361.
44. Friston, K. (2014). The free-energy principle: a rough guide to the brain? *Trends in Cognitive Sciences*, 13(7), 293-301.
45. Ma, J. (2016) Retrieved February 12, 2020, from https://medium.com/@ jianqiangma/all-about-recurrent-neural-networks-9e5ae2936f6e.
46. Maass, W.; Natschläger, T.; and Markram, H. (2002). Real-time computing without stable states: a new framework for neural computation based on perturbations. *Neural Computing*, 14(11), 2531-2560.
47. Boccato, L.; Attux, R.; and Von-Zuben, F.J. (2014). Self-organization and lateral inter action in echo state network reservoirs, *Neurocomputing*, 138, 297-309.
48. Nectoux, P.; Gouriveau, R.; Medjaher, K.; Ramasso, E.; Chebel-Morello, B.; Zerhouni, N.; and Varnier, C. (2012). PRONOSTIA: An experimental platform for bearings accelerated degradation tests. *IEEE International Conference on Prognostics and Health Management*, Colorado, USA, 1-8.
49. Yoo, Y.; and Baek, J.G. (2018). A novel image feature for the remaining useful lifetime prediction of bearings based on continuous wavelet transform and convolutional neural network. *Applied Sciences*, 8(7).
50. Wang, Y.; Peng, Y.; Zi, Y.; Jin, X.; and Tsui, K.L. (2016). A two-stage data-driven based prognostic approach for bearing degradation problem. IEEE *Transactions on Industrial Electronics*, 12 (3), 924-932.