# Comprehensive Predictive Maintenance of Marine Equipment using LSTM Neural Networks

Vineet Gupta* PhD

* *QED Analyticals, INDIA*
* Corresponding Author. Email: vineet@qedanalyticals.com

**Synopsis**

Critical and auxiliary equipment aboard ships is maintained using a combination of preventive and corrective maintenance. These policies lead to over maintenance and unanticipated failures which can be very costly since they compromise operational readiness of ships. Predictive maintenance based on monitoring the current health of the component can reduce unanticipated failures and reduce costs associated with unnecessary maintenance. In this paper, An AI and IoT based comprehensive Predictive Maintenance Framework is proposed for equipment aboard ships, in particular for a water pump. Components which have a high frequency of failure and result in significant downtime are chosen for Predictive Maintenance. Vibration, acoustics, temperature, current, pressure and flow are continuously monitored in real time by installing sensors at multiple pick points on the equipment. A novel single layer Long Short Term Memory (LSTM) neural network model with an attention module is used to predict the degradation level of the components and the most common failures of the pump. The optimal value of the degradation level at which maintenance should be performed for the subsystem is calculated using cost analysis of maintenance. However the failing component is likely to be interdependent on several other components. A cost based grouping model clusters components taking into account the predicted health of the components, the economic dependencies among the components and the location of the components within the system. In the case of a maintenance alert the cluster of components would undergo joint maintenance to optimize maintenance costs.

Keywords: Predictive Maintenance, AI, IoT, Sensors, Neural Network. Marine Machinery

## 1. Introduction: Maintenance Paradigms

Maintenance of critical and auxiliary equipment aboard ships is currently carried out by a combination of preventive and corrective maintenance. Preventive maintenance is a crude policy where the time/usage intervals of scheduled maintenance are determined by the OEM and often leads to unnecessary maintenance. Under corrective maintenance the component is run till failure. Both Preventive and Corrective maintenance lead to unanticipated downtime which leads to operational losses. On Naval ships it can lead to significant costs since it compromises combat preparedness. An AI based predictive maintenance (PdM) paradigm which predicts component failures, based on the current health of the component, could reduce the cost of unnecessary maintenance and unanticipated failures.

In this paper, the author proposes a comprehensive Predictive Maintenance Framework for marine equipment, in particular for a large centrifugal pump with a flow rate of $300m^3$/hr based on monitoring the condition of the pump through IoT sensors. The author presents a method to choose critical components that are suitable for predictive maintenance. Data collected from the sensors is analyzed using a novel Long Shor Term Memory (LSTM) model with an attention module that will predict the degradation level and the Remaining Useful Life (RUL) of the critical components. However, a knowledge of the RUL of a component may not provide the operator complete information about when and which components should undergo maintenance. As part of the post predictive decision-making module, a cost based

---

**Author's Biography**

**Vineet Gupta** is the Founder and CEO of QED Analyticals, India. He holds a BS (Physics, Harvard University), MA (Mathematics, University of Cambridge), PhD (Mathematics, California Institute of Technology). He has been responsible for the development and deployment of several Big Data and AI projects including Predictive Maintenance solutions for government and private organizations.

function determines the exact time when maintenance should be performed to optimize the cost of maintenance.  Furthermore, a cost-based clustering method is used to establish which cluster of components should undergo joint maintenance.  Based on these calculations, an optimal amount of inventory can be stocked.  In our opinion this is the first time a complete predictive maintenance framework from the viewpoint of the operator has been proposed.

## 2. Current Research and Challenges

The ability to predict the need for maintenance of machinery at a specific future moment is one of the biggest challenges for Industry 4.0.  As a result Predictive Maintenance (PdM) has received significant attention in recent times, see (Snchez, 2016) for recent overviews.

Generally, the predicted maintenance framework consists of two components:

1. Prediction of the level of degradation of the machine, including the remaining Useful Life (RUL)
2. Making decisions based on these predictions.

Based on the predictive approaches, most studies can be classified into:

1. Rule based approaches
2. Data Driven approaches

The first approach relies on creating a stochastic model of the degradation evolution of the system (Papakonstantinou, 2014). This approach uses sensors which monitor the machine health and sends alerts based on predefined settings, once a specific rule has been activated. The advantage of this approach is that it does not require large amounts of data relating to machine faults and failures. It also provides simple clear and unambiguous alerts to the operator based on easy to understand rules and thresholds.  However developing such a solution requires extensive prior knowledge of system degradation processes. Furthermore, it is difficult to develop complex models of the deterioration of systems that often consists of thousands of components. From a practical view, such models are rarely able to capture the entire gamut of real life operational variables. As a result these models often make simplifying assumptions that lead to wrong maintenance decisions.

In recent times the data-driven framework has received a lot of attention (Loutas, 2013). These models use large data sets of machine health to predict the health of the system in the future.  The traditional data-driven approaches (Medjaher, 2012) require manual processing and analysis of data by human experts.  As the amount of data being collected has grown exponentially, in recent studies, the deep learning (DL) methods which do not require data wrangling and signal processing techniques have become popular (Deutsch,2018).  However, there are significant challenges to employing deep neural networks for predictive maintenance.

The first is the lack of meaningful data sets.  Deep learning predictive models require the availability of datasets with run-to-failure trajectories of individual components.  This time series data needs to be annotated with proper mode of failure labels.  Such data is rarely available because of the multitude of failure modes, each of which is a rarely occurring event. Even if OEM's possess this data, they are reluctant to share it in the open community.

Currently, most available datasets are synthetic datasets generated with simulators or developed in a lab environment for simple systems. (Eker, 2012)

Both model and data driven methods focus on either diagnosing a fault or predicting the Remaining Useful Life (RUL) of components. They do not estimate the degradation level at which maintenance should be performed. Determining the threshold level of degradation at which maintenance should be carried out requires considering the cost of maintenance and the cost incurred by unanticipated failure.

These models do not determine the type of maintenance to be performed, i.e. which group of components need to be maintained jointly. Components in complex systems are interdependent since the functioning of one component affects the functioning of other components. Maintenance costs can be optimized by clustering interdependent components for joint maintenance. This clustering needs to take into account the degradation level of the components, the interdependency of the components and the physical proximity of the components.

In this paper a hybrid system is proposed which incorporates both Deep Learning and Stochastic modelling to provide a complete predictive maintenance framework. The solution not only predicts the future health of the system but provides clear recommendations on when and what type of maintenance is required.

## 3. Approach

### *3.1. Risk Profile Number*

The suitability of components for predictive maintenance was based on the criticality of the faults, the frequency of the faults and if those faults could be detected. Components were chosen for predictive maintenance by assigning a Risk Profile number based on criticality and frequency of breakdowns. A FMECA was conducted to determine the most critical failures based on the down time associated with the particular failure. This was done using data provided by the pump OEM. The data of the frequency of failure modes was also provided by the OEM. Failures which resulted in a down time of more than 24 hrs. and which had a frequency higher than 1 failure every 5 years were determined as suitable candidates for predictive maintenance. The list of these failures and the associated components is given in Sec 4.

### *3.2. Scaling Factor*

Marine equipment is often customized or manufactured in limited quantities which makes it challenging to obtain data for a particular equipment. The centrifugal pump under consideration was a custom manufactured equipment. A fully functional scaled down version of the pump was obtained from the manufacturer. The specific speed of a pump is defined as:

$$N_s = \frac{N\sqrt{Q}}{H^{3/4}} \tag{1}$$

where:
$N_s$ = specific speed
$N$ =RPM
$Q$ =Flow rate at Best Efficiency Point (BEP)
$H$ =Head at BEP

The specific speed is a useful index to compare different pumps. The specific speed of the scaled down pump was similar to that of the original pump to ensure the degradation behavior was similar up to a scaling factor. The study was carried out on the scaled down version of the pump.

### *3.3. Data Set*

A meaningful data set was created by inducing particular faults in components and letting the components degrade under accelerated load conditions. Sensors placed using noninvasive methods monitored the degradation level of the components. In this way data relating to normal and failure modes was generated for the deep learning model.

### *3.4. Stochastic Model*

The degradation state of the system at any time *t* is modelled as a stochastic process $(\boldsymbol{Y_t})_{t\geq 0}$. $\boldsymbol{Y_t}$ is a vector valued random variable which represents the degradation level of the system at time t:

$$\boldsymbol{Y_t} = \begin{pmatrix} Y_{1t} \\ Y_{2t} \\ \cdot \\ \cdot \\ Y_{rt} \end{pmatrix} \tag{2}$$

where $Y_{it}$ represents the degradation level of the $i$th component at time $t$.

The following assumptions are made related to the degradation of component *i* (*i=1,2…r*):

1. The condition of component *i* at time *t* can be measured and represented by a scalar valued random variable $Y_{it}$. The degradation trajectory is a sample path of the time-dependent stochastic process $(Y_{it})_{t\geq 0}$.
2. $Y_{it}$ has a lognormal distribution with mean $\mu_{it}$ and variance $\sigma_i$. The mean $\mu_{it}$ will be a monotonically increasing function of *t* while the variance $\sigma_i$. Is assumed to be independent of *t*. Lifetime distributions of components in which physical fatigue and stress are the primary contributors of physical failure are often modelled as log normal distributions. This assumption is required to impose additional structure on the data since there are a limited number of sample paths to estimate the probability of failure of a component.
3. The initial degradation level of component *i* is $Y_{i0} = 0$. The component *i* is considered initially as new. The higher the value of $Y_{it}$, the closer component *i* is to failure.
4. A critical degradation threshold $Z_i$ is determined based on economic or technical factors. for component *i*. If $Y_{it} = y_{it}$ exceeds $Z_i$ component *i* is considered have failed. A failed component is required to be maintained immediately.

### *3.5. Predictive Model*

The degradation level at time *t*, depends on the how the fault has progressed in the past. An LSTM neural network is used to model the time dependency. LSTM networks keep track of long and short term dependencies and are well adapted to modelling sequential data. The input data to the LSTM or the input sequence consists of the sensor readings over a fixed time period. The output sequence will be the prediction of the degradation level over a fixed time horizon. The model will predict the degradation level of the component rather than the RUL.

This exploits the large amount of degradation data generated by the fault seeding. The data relating to end of life of the component is limited. The RUL of component *i* can be calculated using the degradation level and the critical degradation threshold $Z_i$.

### *3.6. Threshold Degradation Level*

It is required to determine the threshold $Z_i^{lim}$, at which maintenance of component *i* should be carried out, to minimize the total cost of maintenance during the time period under consideration. $C_{pm}$ is defined as the cost of predictive maintenance and includes both the cost of performing maintenance and the cost of planned downtime. Predictive maintenance will be performed if the degradation level of component $i$, $Y_{it} > Z_i^{lim}$.

$C_{rm}$ is the cost of reactive maintenance and includes the cost of maintenance of the failed component and the cost of unplanned downtime. The cost of unplanned downtime is assumed to be higher than the cost of planned downtime.

Total $C_{pm}$ over a period of time 0,1…T can be defined as:

$$Total\ C_{pm} = C_{pm} \times \sum_{t=0}^{T} I_t \qquad (3)$$

Where $I_t$ is an Indicator variable which takes the value 0 if $Y_{it} < Z_i^{lim}$ and 1 otherwise.

Reactive maintenance is performed in the interval $(t+1, t)$ if the component fails in the interval and predictive maintenance is not performed in the interval. Let $P_{it}$ be the conditional probability that the *ith* component fails at time *t*+1 given that it was working at time *t*.

$$P_{it} = Prob\left(Y_{i,t+1} \geq Z_i \middle| Y_{i,t} < Z_i\right) \qquad (4)$$

Total $C_{rm}$ over a period of time 0,1…T can be defined as

$$Total\ C_{rm} = \sum_{t=0}^{T} C_{rm} \times (1 - I_{t+1}) P_{i,t} \qquad (5)$$

The cost function that needs to be minimized to determine $Z_i^{lim}$ can be defined as:

$$Total\ Cost = Total\ C_{pm} + Total\ C_{rm} \qquad (6)$$

The values of $Y_{i,t}$ and $P_{it}$ are determined by the predictive model as explained in Sec 5. The Total Cost function is minimized to determine the value of $Z_i^{lim}$.

### *3.7. Joint Maintenance*

The Health of component is defined as the probability that component *i* has not failed at time *t*:

$$H_i(t) = Prob\ (Y_{it} < Z_i) \qquad (7)$$

For the time interval $(t+1, t)$, the conditional Health, given that that component $i$, is working at time $t$, is given by:

$$H_i(t+1|t) = Prob\left(Y_{i,t+1} < Z_i \middle| Y_{i,t} < Z_i\right) = 1 - P_{it} \qquad (8)$$

A Path Set of a system is a set of components whose functioning ensures that the system is functioning. A Minimal Path Set is defined a as a path set where the removal of any element results in the set no longer being a path set. We consider the set $\Omega$ of all the MPS of the system.

The Health of the system can be defined in terms of the Minimal Path Sets (MPS) as follows:

$$H(t) = 1 - \prod_{MPS \in \Omega} \{1 - \prod_{i \in MPS} H_i(t)\} \qquad (9)$$

The conditional reliability of the system is defined as:

$$H(t+1|t) = 1 - \prod_{MPS \in \Omega} \{1 - \prod_{i \in MPS} H_i(t+1|t)\} \qquad (10)$$

The idea behind joint maintenance is to find the optimal group of components which will increase the Health of the system by the maximum amount compared to the total cost of predictive maintenance of the group. This will be the group of components *G* that maximizes:

$$\frac{H_G(t+1|t) - H(t+1|t)}{C_{pm}^G} \qquad (11)$$

where:

$H_G(t+1|t)$ is the Health of the system if the group of components *G* undergo preventive maintenance at time *t* and $C_{pm}^G$ is the cost of preventive maintenance of the group of components G, which is defined in Eq. (3). It follows from Eq. (8) and Eq. (11) that the determination of G requires the calculation of the Minimal Path Sets of the system, value of $P_{it}$ and the cost of preventive maintenance of all the components.

## 4. Common Failures

The common pump failures and associated components that were chosen for predictive maintenance were:

### *4.1. Mechanical Failure*

*4.1.1. Mode: Bearing Failure*
*Component: DE Pump Bearing/NDE Pump Bearing/DE Motor Bearing/NDE Motor Bearing*

The bearing is an important part of the centrifugal pump that supports the pump and motor shaft. Bearing failure occur due to poor lubrication, overload, and other reasons. These failures are often associated with an increase in temperature, vibrations in the narrow band, and the kurtosis index.

*4.1.2. Mode: Misalignment Fault*
*Component: Pump Shaft/Impeller/Motor Shaft/Rotor*

A displacement or angular deviation of the pump shaft or motor shaft at the coupling results in misalignment. It causes changes in the vibration signals in the axial and radial directions and increases the vibration amplitude at twice the operating frequency.

*4.1.3. Mode: Imbalance Fault and Feature of rotating parts:*
*Component: Pump Shaft/Impeller/Motor Shaft*

Rotating parts can have unbalanced masses due to manufacturing and assembly defects, and blockages during operations. This results in unbalanced forces which changes the vibration frequency and amplitude.

*4.1.4. Mode: Loose Fault and Feature*
*Component: Foundation/Shaft/Impeller/Seals*

Loose faults are caused by loose foundation or poor fit between components. The resulting time wave form of the vibration signal will show impacting and fixed vibration direction. Looseness caused by poor assembly often has spectral characteristics of superimposed working frequency and high-order harmonics.

*4.1.5. Mode: Mode: Winding Failure:*
*Component: Stator Winding/Rotor Winding*

Winding failures cause thermal deterioration resulting in rise in temperature, current/voltage anomalies and vibration increase.

***4.2. Electrical Failure***

*4.2.1. Mode: Current leakage*
*Component: Shaft/Bearing*

Electrical overload or over-current may be caused by high or low voltage supply or short circuited conductors. It can result in current leakage with current flow through shaft and bearings.

***4.3. Fluid Fault***

*4.3.1. Mode: Abnormal Flow Passage and Feature*
*Component: Inlet strainer/Volute/Impeller*

During operations blockages at the inlet and volute and improper assembly of impeller can reduce pump efficiency. In such cases, the vibration amplitude of the impeller increases with an increase in RPM.

*4.3.2. Mode: Mode: Water Hammer Fault and Feature*
*Component: Valve*

Sudden opening and shutdown can result in a sudden change in the flow of fluid through the pump body causing a shock phenomenon. It can result in a sharp increase and then rapid decay of vibrations and acoustic amplitude.

*4.3.3. Mode: Cavitation Fault and Feature*
*Component: Shaft/Impeller/Bearings*

Cavitation results increased vibration, noise and decrease in efficiency of the pump. The vibration and acoustic features of a cavitation fault are a continuous wide band signal.

## 5. Methodology

### *5.1. Data Collection*

The data consists of the measurements made by the sensors at regular time intervals. These include measurements of:

- Vibration (Tri Axial)
- Temperature
- Flow
- Current/Voltage
- Acoustic emissions

The data consisted of normal and failure mode operations. Components were seeded with a particular flaw and the pump operated till the component failed to obtain failure mode data. The data of the normal and failure modes was aggregated. The aggregated data was split into a training subset and a testing subset using an 80:20 ratio. The training data was used to fit the LSTM model while the testing data was used to evaluate the fit of the model.

### *5.2. Predictive Model*

The sensor data observed at time *t*, is represented by the vector

$$\boldsymbol{X_t} = \begin{pmatrix} X_{1t} \\ X_{2t} \\ \cdot \\ \cdot \\ X_{kt} \end{pmatrix} \qquad (12)$$

where $X_{it}$ is the *i*th sensor observation at time *t*.

The output consists of the level of degradation of the r components which at time t is represented by the vector $\boldsymbol{Y_t}$

$$\boldsymbol{Y_t} = \begin{pmatrix} Y_{1t} \\ Y_{2t} \\ \cdot \\ \cdot \\ Y_{rt} \end{pmatrix} \qquad (13)$$

where $Y_{it}$ is the degradation level of component i at time t which is assumed to be monotonically increasing in every component. The degradation levels are measured physically for fixed points in time $t_1, t_2 \ldots . t_p$. For any point in time between two of these fixed points the degradation level will be a linear interpolation of the degradation level at the fixed points. So, for any *t* such that $t_i < t < t_j$,

$$\boldsymbol{Y_t} = \boldsymbol{Y_{t_i}} + (t - t_i)\left(\boldsymbol{Y_{t_j}} - \boldsymbol{Y_{t_i}}\right) \qquad (14)$$

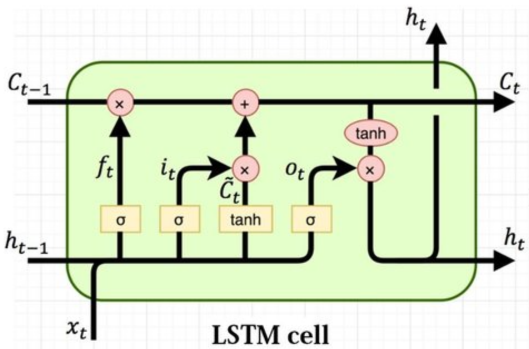
is a piecewise linear function.

It is assumed that

$$\boldsymbol{X_t} = f\,(\boldsymbol{Y_t}) \qquad (15)$$

for some unknown function $f(.)$.

A LSTM neural network model is used to estimate the function $f(.)$. The basic structure of a LSTM cell is given in Fig.1




$$i_t = \sigma\left(x_t U^i + h_{t-1} W^i\right)$$
$$f_t = \sigma\left(x_t U^f + h_{t-1} W^f\right)$$
$$o_t = \sigma\left(x_t U^o + h_{t-1} W^o\right)$$
$$\tilde{C}_t = \tanh\left(x_t U^g + h_{t-1} W^g\right)$$
$$C_t = \sigma\left(f_t * C_{t-1} + i_t * \tilde{C}_t\right)$$
$$h_t = \tanh(C_t) * o_t$$

Fig1: Basic Structure of a LSTM Cell

In an LSTM unit, $C_t$ represents the long term memory ,while $h_t$, the hidden state represents the short term memory at time *t*. There are three gates in a LSTM unit: the forget gate, the input gate, and the output gate. Each of them controls the amount of information used in the unit. Generally, based on the input $x_t$ and $h_{t-1}$, the input gate $i_t$, $\in \mathrm{R}^{k\times 1}$ decides which values to use in calculating $C_t$. Next, forget gate $f_t$, $\in \mathrm{R}^{k\times 1}$ determines which information from $C_{t-1}$should be removed, and which can be used to update $C_t$. Finally, output gate $o_t \in \mathrm{R}^{k\times 1}$ controls what information in $C_t$ to become $h_t$.

A Sequence to Sequence LSTM model is used which will take a sequence of observations and output another sequence. The Input sequence will be $\boldsymbol{X}_t, \boldsymbol{X}_{t+1} \ldots . . \boldsymbol{X}_{t+m}$ and the output sequence will be $\boldsymbol{Y}_t, \boldsymbol{Y}_{t+1} \ldots . . \boldsymbol{Y}_{t+n}$ where *n* > *m*.

A Two layered Encoder-decoder is the standard modelling paradigm for sequence-to-sequence tasks in LSTM. The first layer is the autoencoder layer which converts the input sequence to a single context vector of fixed dimension. The second layer is the decoder layer which converts the context vector into the output sequence as shown in Fig.2. However, there are two major problems with this architecture:

1. Loss of information about higher level relations between multiple sensor readings. For instance, a bearing fault may be picked up early by an ultrasonic sensor and as the fault progresses the vibration and temperature signatures also change.
2. The decoder may require different information at different times which is not possible with the context vector which is the last hidden state vector

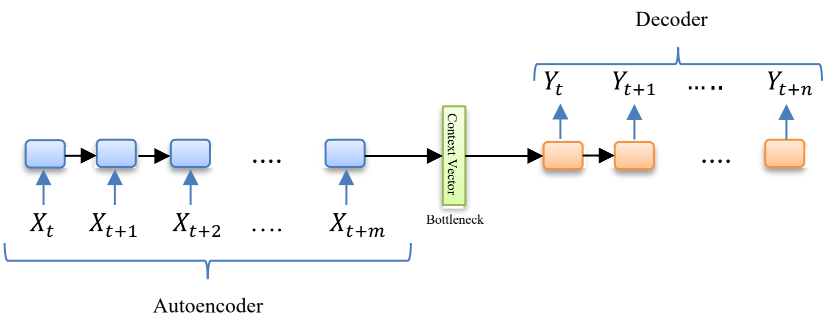



Fig 2: Architecture of Sequence to Sequence LSTM Network

So, a single layer LSTM is employed for a seq to seq model. The architecture is shown in Fig 3.

The single layer Seq to Seq LSTM architecture does away with the bottleneck of the context vector allowing the network to model the high level relations between the sensor measurements.

The standard LSTM model uses only the last hidden state as the output. In the proposed model, an adaptation of the attention model developed by Bhadanau (Bhadanau, 2015) is used. A SoftMax function is applied to all the hidden states which gives a weight to each hidden state. The final output of the attention module is the weighted sum of all the hidden states as shown in Fig 3. This will allow the model to weight the sensor readings with more important readings getting weighted more in accordance with the fault. For instance, a shaft misalignment fault will weight the vibration measurements more than the temperature measurements.

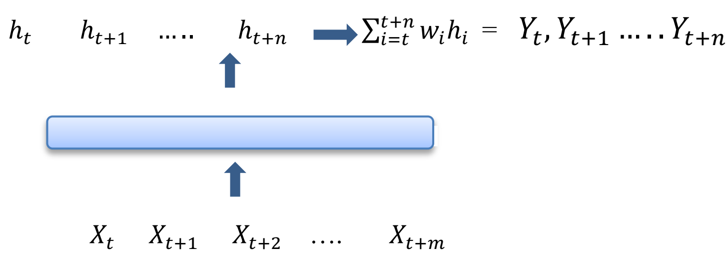



Fig 3: Architecture of proposed LSTM Network

Thresholds for failure of component *i* denoted by $Z_i$, are determined as described in Sec 3.

At time *t,* if the output sequence contains $Y_{it^*}$where $Y_{it^*} = y_{it^*} > Z_i$ , the RUL of the component *i*, is given by:

$$RUL(i) = \min_{t^*}(t^* - t) \qquad where\ y_{it^*} > Z_i \tag{16}$$

As described in Sec 3,

$$logY_{it} \sim N(\mu_{it}, \sigma_i) \tag{17}$$

The parameters of the distribution can be estimated from the data and the value $P_{it}$ is calculated from the log normal distribution. However, lack of sufficient data prevented us from completing this step.

**6. Result**

The sensor data was resampled at 1 minute intervals and the mean value of the data was considered for each interval. A sliding window of size 20 with a step size of 1 was used to generate the input sequences. The output sequence of the model consisted of the predicted degradation level of the components. The Root Mean Square Error (RMSE) was calculated for the predicted values and the observed values for both the test and training data set. The number of training epochs was varied and the RMSE of both the test and training data sets were compared across the epochs. The optimal number of epochs was determined to be 1000, since a higher number indicated overfitting of the model as the RMSE of the training set started to increase. A similar process was used to determine the number of neurons and batch size. The final configuration used in the LSTM was 2 neurons, batch size 8 and trained for 1000 epochs. The RMSE calculated for the aggregated data, which represents the overall accuracy of the model was 97.6%.

The test data was stratified according to failure modes discussed in Sec 4. The accuracy of the model was tested for prediction of each failure mode. The results are given in Table 1.

Table 1: Failure Mode Accuracy Test Results

| **Failure Mode** | **RMSE** |
|---|---|
| Bearing Failure | 97.79% |
| Misalignment Fault | 93.56% |
| Imbalance Fault | 90.79% |
| Loose Faults | 95.58% |
| Winding Failures | 96.67% |
| Current Leakage | 100% |
| Abnormal Flow | 93.29% |
| Cavitation Failure | 92.67% |

Due to lack of data the conditional probability of the failure of a particular component given in Eq.(4) could not be determined from the data.

## 7. Conclusion and Future Work

In this study the author has proposed a comprehensive solution for predictive maintenance of machines, in particular a centrifugal pump aboard a ship. The framework was tested on a scaled down pump with similar specific speed and accurately predicted the most common modes of failure. The model predicted the degradation level of the components rather than the RUL directly. This approach was preferred since there was sufficient data relating to the degradation level of the components, the data relating to failure was limited. The author conjectures the model can be applied to the original pump by comparing the normal signatures of the two pumps and determining a scaling factor. The framework also provides a method to determine the optimal time and the optimal cluster of components that should undergo joint maintenance based on economic costs. At this point lack of sufficient data prevented calculation of the conditional probability of failure of the components as defined in Eq. (4), from the data. This parameter is necessary to determine the optimal time and cost based grouping for optimal joint maintenance in the proposed model as described in Sec 3.6 and Sec 3.7. As the study is still in progress the author expects to obtain the required failure data to validate the model. Validation of the model also requires details of the operating and maintenance costs to determine the total cost of maintenance and the cost of preventive maintenance as defined in Eq. (3) and Eq. (6). These costs will be specific to each user and will have to be obtained from a particular user of the equipment.

## References


Bahdanau D, K. Cho, and Y. Bengio. Neural machine translation by jointly learning to align and translate. In Proc. of the 3rd ICLR, 2015. arXiv:1409.0473.

Deutsch J, He D. Using deep learning-based approach to predict remaining useful life of rotating components. IEEE Trans Syst, Man, Cybern 2018;48(1):11–20.

Eker, Ö.F.; Camci, F.; Jennions, I.K. Major challenges in prognostics: study on benchmarking prognostic datasets. In Proceedings of the 1st European Conference of the Prognostics and Health Management Society 2012, Dresden, German, 3–6 July 2012; Volume 3, p. 8.

Loutas TH, Roulias D, Georgoulas G. Remaining useful life estimation in rolling bearings utilizing data-driven probabilistic e-support vectors regression. IEEE Trans Reliab 2013;62(4):821–32.

Medjaher K, Tobon-Mejia DA, Zerhouni N. Remaining useful life estimation of critical components with application to bearings. IEEE Trans Reliab 2012;61(2):292–302.

Papakonstantinou K, Shinozuka M. Planning structural inspection and maintenance policies via dynamic programming and markov processes. part i: theory. Reliab Eng Syst Saf 2014;130:202–13.

Snchez-Silva Mauricio, Frangopol Dan M., Padgett Jamie, Soliman Mohamed. Maintenance and operation of infrastructure systems: review. J Struct Eng 2016;142(9)