International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-08, Issue-04, July 2022

Prognosis of Rolling Element Bearing Using LSTM Neural Network

Shubham Kate[1], Sharad Gaikwad[2], Ajeet Patil[3]

[1]M. Tech Student, Mechanical Engg. Department, Walchand College of Engineering Sangli.

[2]Assistant Professor, Mechanical Engg. Department, Walchand College of Engineering Sangli.

[3]Assistant Professor, Mechanical Engg. Department, Walchand College of Engineering Sangli.

Abstract: Rolling element bearing is crucial component in mechanical system. It is subjected to challenging conditions which leads to rapid wear and tear in bearing. If such deterioration is not addressed, the overall system has risk of failure or breakdown. Prognostics uses prediction techniques to estimate the failure point of a system. RUL prediction techniques can be classified into two categories. Physics-based methods estimate the remaining time based on the spread of the damage process or physics of failure. Data-driven models take historical failure data from the machine. It helps to predict next system malfunction by building machine learning model. For RUL prediction, numerous efficient data-driven techniques have been introduced. In this work, LSTM neural network was deployed to predict the failure time of bearing. This work is done on FEMTO bearing data set from NASA Ames Prognostic Repository. The statistical features were extracted from raw dataset for construct health indicator. The features and health indicator given as input for training the LSTM model.

***Keywords* — *Prognosis, RUL, Machine learning, LSTM neural network, Regression*.**

I. INTRODUCTION

Important parts of rotating machinery include rolling element bearings. Since these bearings typically function in challenging circumstances, performance degradation is inevitable. If such deterioration is not addressed, the overall system risk of failure or breakdown. Prognostics uses prediction techniques to estimate the Remaining Useful Life (RUL) of a system. It uses historical data, present usage, and anticipated operating circumstances. Event logs and historical failure data from comparable systems are examples of historical information. A characteristic obtained from sensor data that shows the system's present state of health could be used to describe current usage. Future operating conditions are operational and environmental aspects that might have an impact on the system's status in the future. These details can be discovered by consulting industry experts or by looking at the manufacturing schedules. Prognostics has a number of advantages, such as preventing abrupt failure as well as unplanned maintenance. Lengthening component's span of life by informing maintenance personnel in advance of the severity of the problem.

Numerous studies have been conducted in recent years to create techniques for rotating machinery, particularly bearings, to check their condition. In recent years, there have been numerous successful tools, models, and applications as a result of the growing interest in machinery prognostics.

II. LITERATURE

Sham Kulkarni et. al. [1] work places a strong emphasis on comparison between vibration metrics. They are employed in the identification of bearing flaws, to characterize the dispersed defects in the bearing. With simulated defects on the bearing parts, these parameters are measured under various loads and speeds. The sensitivity of kurtosis in monitoring ball bearing condition is also demonstrated in this work. Sangram Patil et al. [2] illustrates the usage of various regression algorithms. For RUL prediction using time-domain information, Random Forest and Gradient Boosting models are used. The features are taken from the vibration signals that are provided. Extracted features are then ranked. Decision tree technique is used for ranking. Model is then created using these features. This models' performance is further confirmed by showing the appropriate learning curves, and hyper-parameters are tweaked via an extensive parameter search. Xiaochuan Li et. al. [3] proposed a combination of two machine learning techniques. The RUL of bearings was forecast using this combination. Regression model and multilayer artificial neural network model were combined. The bearing failure stages were identified using an analysis of root mean square and kurtosis. Two case studies were used to validate the suggested methodology. Apakrita Tayade et al. [4] calculated the Remaining Useful

 DOI : 10.35291/2454-9150.2022.0402

Life (RUL) for a deteriorated bearing using vibration data produced using a machine learning approach. Regression models and statistical feature extraction analysis are used in the process. Feature selection is dine using principal component analysis. Regression models use an input parameter generated by Principal Component Analysis (PCA) for the training and testing of models. Support Vector Regressor and Random Forest are the two models utilized. The RUL of the bearing is computed using the Weibull Hazard Rate Function. Pangun Park et. al. [5] merged analytical model-based and data-driven methodologies and established the RUL prediction method. A reliable and predictable health indicator is created from feature importance ranking and principal component analysis of variety of features. The degradation model's parameters are then refined using the adaptive sliding window method. This is based on ridge regression of a time series sequence. When compared to conventional Bayesian techniques, the suggested adaptive system significantly outperforms them in terms of RUL estimate accuracy against potential degradation model defects. Patrick Nectoux et. al. [6] introduced an experimental platform named PRONOSTIA. It allows testing, verification, and validation of bearing health assessment. It also helps in assessment of diagnostic, and prognostic approaches. Bearings regarded as crucial because their failure drastically reduces the availability and security of equipment. Since these parts are responsible for the majority of rotating machine problems, selection of bearings is justified. Tarek Berghout et. al. [7] provided summary of latest advances in RUL prediction. Recent Machine learning (ML) methods for RUL prediction in various critical systems were examined. The explanation made sure that the RUL prediction process is simple. Additionally, their study offers detailed instructions for choosing the best approach for any particular driven data type. This manual is followed by a classification of various ML tool types to address all the instances mentioned. These principles are ultimately used in this review-based analysis to identify learning model constraints, reconstruction difficulties, and potential future applications. Tarek Berghout et. al. [8] developed sequence-by-sequence deep learning (DL) algorithm. By using the knowledge, it has learned from the life cycles of other similar systems, it can increase its ability for generalization. The new approach uses a Long-Short-Term Memory (LSTM) neural network as the main building block of adaptive learning to assess health condition. Both health stage and health index estimations are extracted. From experimental validation implies that deep learning knowledge transfer-based prognosis strategy is capable and performs better than other prognosis methods.

III. LSTM FOR PROGNOSIS

Recurrent neural network is used to solve problems involving sequence prediction. Because they carry out the same task for each element of a sequence while relying on earlier calculations, RNNs are known as recurrent neural networks. LSTM is deployed for time-series data prediction and forecasting. An LSTM network is a type of RNN. It can discover long-term relationships between sequence data's time steps.

Recurrent neural network is used to solve problems involving sequence prediction. Because they carry out the same task for each element of a sequence while relying on earlier calculations, RNNs are known as recurrent neural networks. LSTM is deployed for time-series data prediction and forecasting.

The construction of a straightforward LSTM network for regression is shown in Fig. 1. Sequential input layer is first layer. It is followed by LSTM layer. A regression layer is output layer. The connected layer sits between LSTM and regression layer.

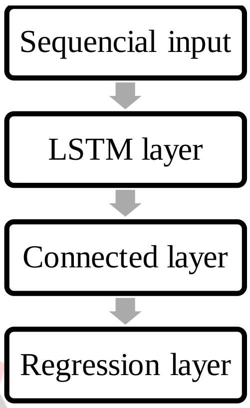

Fig. 1: LSTM architecture

Sequential Input Layer is a first layer in this model. It inputs sequence data to a network. LSTM Layer is second layer in network. It takes input from sequential input layer. Fig. 2 shows structure of LSTM cell. Function of cell is to retain information. The gate is utilized for memory manipulation. There are three gated in LSTM. The forget gate purges the data that is no longer relevant in the cell state. The gate receives two inputs, x (t), input at the current time and h (t-1), prior cell output. They are multiplied with weight matrices before bias is added. The output of the activation function, binary. If a cell state's output is 0, the piece of information is lost. If a cell state's output is 1, the information is saved for use in the future.

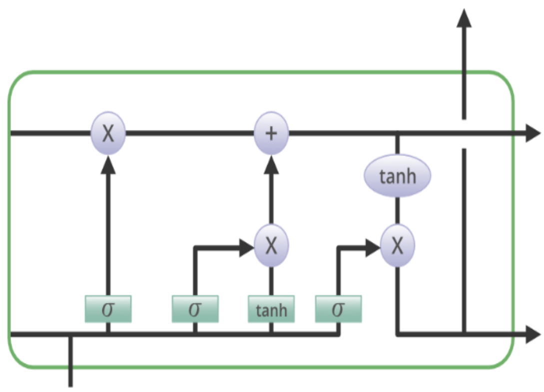

Fig. 2: Structure of LSTM cell

 DOI : 10.35291/2454-9150.2022.0402

ISSN : 2454-9150 Vol-08, Issue-04, July 2022

The input gate updates the cell state with relevant details. The inputs h(t-1) and x (t) are used to regulate the information using the sigmoid function. It filters the values to be remembered similarly to the forget gate. The tanh function is then used to construct a vector with a value between -1 and +1. It includes each and every conceivable value for h (t-1) and x. (t). To extract the valuable knowledge, the vector's values and the controlled values are finally multiplied. The output gate's job is to take meaningful information out of the current cell state and deliver it as output. The tanh function is first used to the cell to create a vector. The sigmoid function is then used to control the information. Using the inputs h t-1 and x t, the data is then filtered by the values to be remembered. The vector's values and the controlled values are finally multiplied and supplied as input and output to the following cell, respectively.

The third layer of an LSTM network is the connected layer. It adds a bias vector after multiplying the input by a weight matrix. Some ML algorithms, particularly neural networks, have learnable parameters called weights and biases, or 'w' and 'b'. A neural network's fundamental building blocks are neurons. Every neuron in a layer of an Artificial Neural Network (ANN) is linked to some or all of the neurons in the layer above it. The weights and the bias are added to the information when they are passed across neurons. Weights regulate the magnitude of the relationship between two neurons. A weight establishes the degree to which an input will affect an output. Biases are an extra input into the following layer. Bias is not modified by the prior layer. It is a constant. However, they do have outgoing connections with their own weights as they lack any incoming connections. The bias unit ensures that the neuron will still be activated even if all the inputs are zeros. The fourth and output layer of this model is regression. Regression is a technique for figuring out how independent features relate to a dependent feature. Once the link between the independent and dependent variables has been estimated, outcomes may then be predicted. Regression is a statistical study area that is essential to machine learning forecast models. It is used as a method to forecast continuous outcomes in predictive modelling. Regression using machine learning often entails drawing a best fitting line through the observations. To obtain the best fit line, the distance between each point and the line is minimized.

IV. EXPERIMENTAL STUDY

In order to test and verify bearing's fault detection, diagnostic and prognostic methodologies, PRONOSTIA setup is developed (Fig. 3). The FEMTO Institute was responsible for designing and implementing the platform. The primary goal of the setup is to deliver accurate experimental data that characterizes the degeneration of ball bearings during their entire operating life. The data given by setup is consistent with bearings that are deteriorated naturally over the time. This demonstrates that the faults start out on the bearings after starting the experiment. Each worn-out bearing has practically every form of flaw. Furthermore, the current setup design permits to offer data relating to bearings deteriorating under varying operating conditions.

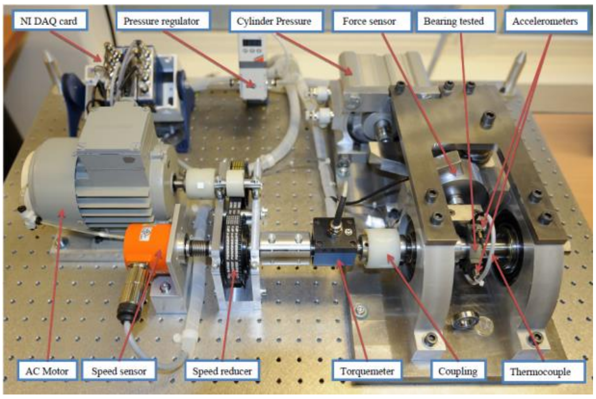

Fig. 3: Experimental setup

Run-to-failure experiment are performed on test setup. Tests were terminated when the vibration signal's intensity exceeded 20g. It is to prevent damage from spreading over the entire test bed. Total seventeen bearings were tested under three different conditions. Table 1 describes the conditions in which bearings were tested. Six bearings allotted for training the model and remaining eleven for testing the model. Table 2 shows this classification of bearings.

Table 1: Conditions for testing of bearing

	Condition 1	Condition 2	Condition 3
Speed (RPM)	1800	1650	1500
Load (N)	4000	4200	5000

Table 2: Dataset classification

Datasets	Operating Conditions		
	First condition	Second Condition	Third condition
Learning Datasets	Bearing 1_1	Bearing 2_1	Bearing 3_1
	Bearing 1_2	Bearing 2_2	Bearing 3_2
Testing Datasets	Bearing 1_3	Bearing 2_3	Bearing 3_3
	Bearing 1_4	Bearing 2_4	
	Bearing 1_5	Bearing 2_5	
	Bearing 1_6	Bearing 2_6	
	Bearing 1_7	Bearing 2_7	

V. MODEL CONSTRUCTION

Fig. 4 shows the steps involved in LSTM model training and testing.

 DOI : 10.35291/2454-9150.2022.0402

International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-08, Issue-04, July 2022

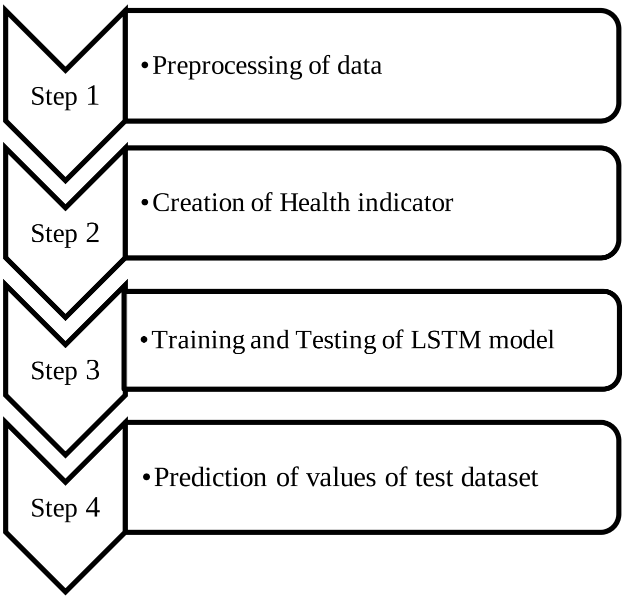

Fig. 4: Steps involved in LSTM model construction

A. *Preprocessing*

An individual measurable quality or characteristic of an event being studied can be defined as a machine learning data feature. First step in the process is turning unprocessed data into numerical variable. Then it will be given as input while preserving the data in the original dataset. This process is feature extraction. Compared to directly using machine learning on raw data, it produces better outcomes. Techniques for feature extraction can also result in benefits including increased accuracy, a lower risk of overfitting, a faster learning curve, and better data visualization. Table 3 shows the extracted features from raw dataset. This feature then given as input for training the LSTM model individually.

Table 3: Features extracted from raw data

KURTOSIS	$N \times \frac{\sum_i^N (Y_i - \bar{Y})^4}{\sum_i^N (Y_i - \overline{Y^2})^2}$
MEAN	$\frac{1}{n}\sum_{i=1}^{n} X_i$
Root Mean Square (RMS)	$\sqrt{\frac{1}{N} \times (x_1^2 + x_2^2 + .. + x_N^2)}$
SKEWNESS	$\frac{\sum_i^N (X_i - \bar{X})^3}{(N-1) \times \sigma^3}$
STANDARD DEVIATION (STD DEV)	$\sqrt{\frac{\sum (X - \mu^2)}{N}}$
VARAINCE	$\frac{\sum (X - \mu^2)}{N}$

B. *Construction of health indicator*

The best representation of the Health Indicator is an exponential function. It represents the degradation of bearing. This is because the method relies on data that are produced from accelerated degradations. The degradation of bearing increases exponentially as time passes. The similarity between the shape of the deterioration and the exceptional function of degradation creates a kind of compatibility between the labels and the extracted patterns. d and a can be obtained using equation (1) and (2)[7][8].

$$HI(t) = d - e^{at+b} \quad \dots (1)$$

t - time instant,

a, b, d - hyperparameters that control the divergence characteristics of the exponential degradation. These parameters are tuned according to the best results of the loss function.

$$\begin{cases} HI\,(tmin) = 1 \quad Normal\ bearing \\ HI\,(tmax) = 0 \quad failed\ bearing \end{cases} \quad \dots (2)$$

C. *Training and testing of LSTM model*

Table 4: Parameters setup for LSTM model

Parameters	**Values**
Hidden layers	20
Epochs	200
State activation function	Tanh
Gate activation function	Sigmoid
Initial learning rate	0.01

Table 4 shows the parameter setup for LSTM model. Fig. 5 shows the output from regression layer of model.

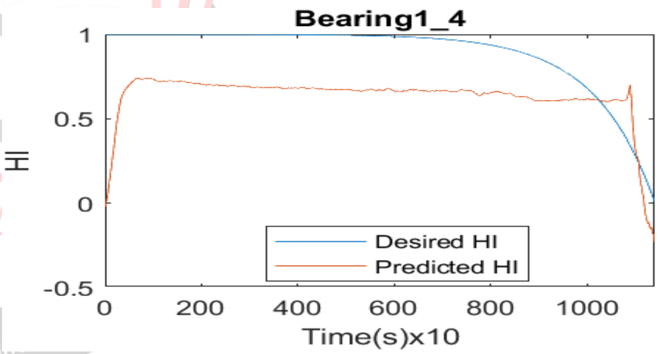

Fig. 5: Curve fitting of prediction of LSTM with variance as input

D. *Model Evaluation*

a) Root Mean Square Error (RMSE)

The precise fit of the model to the data, or how closely the actual data points match the values anticipated by the model, is shown by RMSE. Root Mean Square Error is the standard deviation of the residuals or prediction errors. Residuals are a measure of how far from the regression line data points are; RMSE is a measure of how spread out these residuals are. In other words, it tells you how concentrated the data is around the line of best fit.

b) Score

This metric is provided for this specific challenge. It is calculated by using equations (4), (5) and (6).

 DOI : 10.35291/2454-9150.2022.0402

$$Score = \frac{1}{11} \times \sum_{i=1}^{11} A_i \qquad \ldots (4)$$

$$A_i = \begin{cases} exp^{-\ln(0.5)\times\frac{error}{5}} & if\ error \leq 0 \\ exp^{\ln(0.5)\times\frac{error}{20}} & if\ error > 0 \end{cases} \qquad \ldots (5)$$

$$\%\,Error = 100 \times \frac{Actual\ RUL - \widehat{RUL}}{Actual\ RUL} \qquad \ldots (6)$$

VI. RESULT AND DISCUSSION

Table 5 shows LSTM neural network' evaluation parameters. LSTM Model is judged based on two parameters, Model score and RMSE.

Table 5: LSTM model evaluation

Input Feature	Training RMSE	Testing RMSE	Model Score
KURTOSIS	0.01915	0.3141	0.4486
MEAN	0.2094	0.2170	0.6181
RMS	0.2092	0.2981	0.5124
SKEWNESS	0.2051	0.2820	0.4073
STD. DEV.	0.1866	0.2927	0.4859
VARAINCE	0.2373	0.3693	0.3801

The precise fit of the model to the data, or how closely the actual data points match the values anticipated by the model, is shown by RMSE. Better fit is shown by smaller RMSE values. RMSE is a useful metric for evaluating how well the model estimates the outcome. For primary goal of the model is prediction, then this fit criterion is crucial.

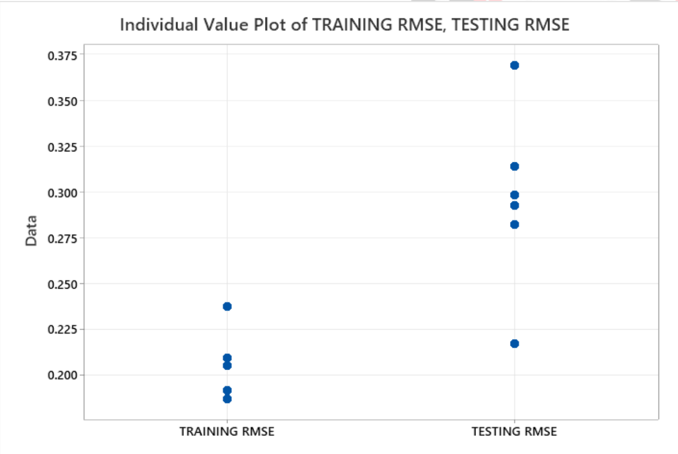

Fig. 6: Individual value plot of training and testing RMSE

Fig. 6 shows individual value plot of training RMSE and testing RMSE. Mean and standard deviation of training RMSE is 0.2065 and 0.0178 respectively. This shows RMSE of different input feature in training is close to each other. Mean and standard deviation of training RMSE is 0.2955 and 0.0493 respectively. This shows RMSE of testing is spread out compared to training RMSE of different input features.

Table 6: Best input feature for LSTM model

INDEX	INPUT FEATURE	MODEL SCORE
1	VARAINCE	0.3801
2	SKEWNESS	0.4073
3	KURTOSIS	0.4486
4	STANDARD DEVIATION	0.4859
5	RMS	0.5124
6	MEAN	0.6181

Model score is another evaluation parameter for LSTM neural network. This parameter specifically designed for this model building problem. It evaluates model based on percentage error between predicted output and actual output. The score is calculated by equation (4). The lower the model score better the model. Table 6 shows best input feature for LSTM model-based model score evaluation parameter. LSTM model with skewness as input feature is best model.

VII. CONCLUSION

This research covers rolling element bearing characteristics in time domain. The extracted time domain features represent the degradation of bearing accurately. In this work, a neural network was utilized to predict the failure time of bearing. This work is done on FEMTO bearing data set from NASA Ames Prognostic Repository. Following conclusions can be drawn from this work:

A. Six statistical features i.e., standard deviation, root mean square, kurtosis, variance, skewness and mean are extracted from raw data. It helps in reducing the sample size of input dataset.

B. RNN based LSTM model is established on training dataset. The model fits the input data correctly as mean RMSE of training and testing is 0.2065 and 0.2955 respectively.

C. Variance is best input feature for LSTM model on basis of model score out of six tested features. It is followed by skewness, kurtosis, standard deviation, root mean square and mean.

VIII. REFERENCES

[1] Patrick Nectoux, Rafael Gouriveau, Kamal Medjaher, Emmanuel Ramasso, Brigitte Morello, Noureddine Zerhouni, Christophe Varnier, "PRONOSTIA: An Experimental Platform for Bearings Accelerated Life Test", *IEEE International Conference on Prognostics and Health Management*, Denver, CO, USA, 2012.

[2] Sham Kulkarni and Anand Bewoor, "Vibration based condition assessment of ball bearing with distributed defects", Journal of Measurements in Engineering, Vol. 4, Issue 2, 2016, p. 87-94.

[3] Sangram Patil, Aum Patil, Vishwadip Handikherkar, Sumit Desai, Vikas Phalle and Farukh Kazi, "Remaining Useful Life (RUL) Prediction of Rolling Element Bearing Using Random Forest and Gradient Boosting Technique", Proceedings of the ASME 2018- International Mechanical Engineering Congress and Exposition, November 9-15, 2018, Pittsburgh,

 DOI : 10.35291/2454-9150.2022.0402

PA, USA; DOI: https://doi.org/10.1115/IMECE2018-87623.

[4] Xiaochuan Li, Faris Elasha, Suliman Shanbr and David Mba, "Remaining Useful Life Prediction of Rolling Element Bearings Using Supervised Machine Learning", Energies, Vol 12, Issue 14: 2705, 2019; DOI: https://doi.org/10.3390/en12142705.

[5] Apakrita Tayade, Sangram Patil, Vikas Phalle, Faruk Kazi and Satvasheel Powar, "Remaining useful life (RUL) prediction of bearing by using regression model and principal component analysis (PCA) technique", Vibroengineering PROCEDIA, Vol. 23, 2019, p. 30-36; DOI: https://doi.org/10.21595/vp.2019.20617.

[6] Pangun Park, Mingyu Jung and Piergiuseppe Di Marco, "Remaining Useful Life Estimation of Bearings Using Data-Driven Ridge Regression", Applied Sciences, Vol 10, Issue 24: 8977, 2020; DOI: https://doi.org/10.3390/app10248977.

[7] Berghout, Tarek, and Mohamed Benbouzid. 2022. "A Systematic Guide for Predicting Remaining Useful Life with Machine Learning", Electronics 11, no. 7: 1125. DOI: https://doi.org/10.3390/electronics11071125.

[8] T. Berghout, L. -H. Mouss, T. Bentrcia and M. Benbouzid, "A Semi-supervised Deep Transfer Learning Approach for Rolling-Element Bearing Remaining Useful Life Prediction," IEEE Transactions on Energy Conversion, DOI: https://doi.org/10.1109/TEC.2021.3116423.

[9] "FEMTO Bearing Data Set", NASA Ames Prognostics Data Repository, http://ti.arc.nasa.gov/project/prognostic-data-repository, NASA Ames Research Centre, Moffett Field, CA.

