EECS 545 Final Report:
Predicting Rolling Bearing Remaining Useful of Life based on Data-driven Methods

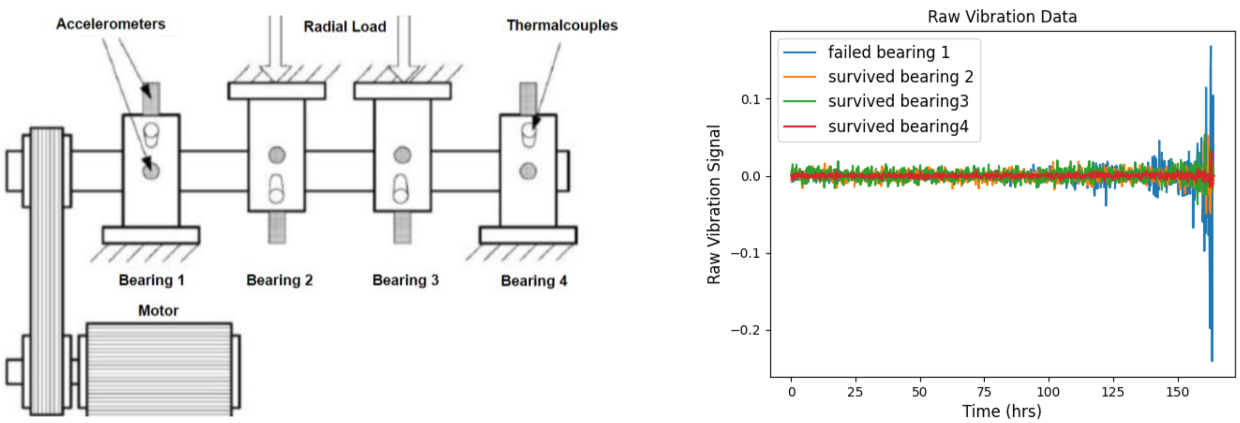
Yizhou Chen, Tianyu Shou, Wenxi Li, Chuan He, Bingqian Yang
{yizhouch, tshou, wenxli, chuanum, bingqian}@umich.edu

Abstract: Rolling bearings are one of the most critical components of rotating machinery. Any fault or failure can lead to poor fabrication performance, failure of the entire system, even catastrophic accidents. Accurate prediction of rolling bearings degradation during long-time operation under extreme harsh environments is of great importance. In this project, a novel intelligent remaining useful life (RUL) prediction method - Two-stage degradation model is implemented on a natural run-to-failure bearing datasets from Center for Intelligent Maintenance Systems (IMS). The degradation data is separated into slight and severe degradation stages and fitted into linear regression model and exponential random coefficient model, respectively. The accurate RUL of bearing is predicted by our multi-state degradation model within an acceptable time interval. Based on the experiment results, the two-stage degradation model has superior performance in predicting the RUL of rolling bearings compared to RNN+LSTM and CNN+LSTM frameworks. This study will be useful in forecasting the fault status of the rolling bearings and detecting the initial stage of failure for general cases in industrial applications.

1. Introduction

Rotating machinery, a high-speed rotating production equipment, plays a crucial role in manufacturing, aerospace, metallurgy, and the military industries. Rolling bearings as one of the most important load-bearing components in such machines always operate under extremely harsh environments such as extreme rotating speed, high ambient temperature, moisture, and overload [1]. Any misalignment, fault or damage of rolling bearings can result in equipment shut down, high maintenance cost, and even serious safety incidents. Therefore, accurate remaining life prediction (RUL) and defect detection of rotating bearings is the key to guarantee machine system operation safety and reliability.

Accurately identifying the bearing faults is a non-trivial task in practice, especially in the initial failure stage. Currently, RUL prediction approach is based on conventional physical-model based analysis. It is constructed based on mechanical parameters such as working condition, material fatigue strength, and failure criteria [2,3]. However, a physical-based model is typically developed for a specific equipment or system and the cost of the model construction process is expensive. In addition, the accuracy of analysis is always heavily affected by the signal noise and working conditions [4,5]. Thanks to the rapid development of artificial intelligence, data-driven methods based on statistical theory which reflect the realistic dynamic behavior of rolling bearings have the ability to overcome above shortcomings. In existing research, people have implemented classic ML algorithms considerably, such as in artificial neural networks [6], principal component analysis [7], K-nearest neighbors [8]. In recent years, deep learning methods have emerged as one of the most effective solutions for RUL problems. In the deep learning fields, such as convolutional neural networks [9], auto-encoders [10], deep belief networks [11] and recurrent neural networks [12] have been around for years. Recently, an innovative framework combined with CNN and long short-term memory (LSTM) is applied to bearing diagnoses and exhibit promoted prediction performance [13]. The model is implemented with hyperparameter tuning in this project. However, the outcome does not show good performance due to the scarcity and abnormality of the dataset. Based on the knowledge of subject matter expertise [17] and data failure trend, the bearing RMS signal is divided into slight and severe degradation stages [16]. The proposed two-stage degradation model is more suitable to capture the failure features and give better performance. The detailed implementation will be discussed in the proposed method section.



(a) Experimental setup (b) Raw vibration signals

Figure 1. Experimental setup of bearing test rig and raw vibration signals.

2. Dataset description:

A proper dataset is essential for effective machine learning algorithms. Most popular rolling bearing datasets such as CWRU dataset, Paderborn University datasets, and Pronostia datasets adopted accelerated life testing methods (i.e. artificially induced scratch or drill to the bearing surface). Different from these datasets, intelligent maintenance systems datasets (IMS) generated by the NSFI/UCR Center records a complete natural degradation process makes it especially suitable for predicting RUL [13]. A special test rig is designed to apply constant load and record the entire run-to-failure signal of bearings. A constant rotation speed 2000 rpm and a radial load of 6000 lbs. are exerted to the system. Raw vibration data was recorded every ten minutes for a duration of one second, as shown in Figure 1a, then denoised and preprocessed with RMS. Detailed data description can be found on NASA's Prognostics Data Repository. Three test-to-failure experiments are run individually, where the raw vibration data of one trial is shown in Figure 1b. Each trial consists of 20,480 points with the sampling rate set at 20 kHz.

3. Proposed Method

3.1 Feature Extraction

A proper feature plays an essential role of a health indicator to successfully detect the onset of degradation stage, distinguish health and failure state of bearings, and represent prominent degradation trends. There are numbers of features that exist in the domain of time or frequency domain. For instance, mean, peak to peak and root of mean square (RMS) in the time domain. In addition, peak frequency, mean frequency, and amplitude frequency are commonly used in published literature. The current work shows that extracting features from the raw data in the time domain is more suitable for real-time RUL prediction than in the frequency domain [15]. The raw data is very sensitive to external perturbations, such as unexpected vibrations, moisture, and temperature which corrupt the data acquisition process and conceal real system response. To avoid the above deficiencies, the root mean square (RMS) extracted from the vibration signals in the time domain is selected in this project because it is regarded as one of the most suitable and stable statistical features for prognosis analysis. Due to the scarcity of dataset, other features such as mean and peak to peak are also explored to check their feasibility.

3.2 Deep Neural Network

Two machine learning strategies are applied but their performances are not doing well in RUL prediction. RNN/LSTM approach cannot capture the severe failure ascending trend due to the limited training data points, and sensitivity to noise perturbation [18]. To solve this problem, a deeper neural network, which is convolutional and stacked bidirectional/unidirectional LSTM network (SBULSTM), is introduced. As shown in Table.1, it uses CNN layers to nest all the past local information, and then pass information to SBULSTM for information storage and integration for time-series prediction. And finally use linear regression layers to summarize all the information and predict the final model. The dropout layers are used frequently to avoid overfitting. Even though this method's performance has been proved in the paper [19], it is not applicable in our case because the model is too deep for our limited training dataset. After iterative hyperparameter tuning, neither of these models works. Therefore, a two-stage degradation model is introduced.

Table 1: Layer details of the proposed model (channels=18, sequence length=500)

CNN-SBULSTM	Layer number	Description	Details
CNN	Layer 1	1D-convolution	filters = 36, kernel_size = 10, strides = 1, padding = same, activation = *ReLu*
		1D-max-pooling	pool_size = 2, padding = same
	Layer 2	1D-convolution	filters = 72, kernel_size = 10, strides = 1, padding = same, activation = *ReLu*
		1D-max-pooling	pool_size = 2, padding = same
	Layer 3	1D-convolution	filters = 114, kernel_size = 10, strides = 1, padding = same, activation = *ReLu*
		1D-max-pooling	pool_size = 2, padding = same
	Layer 4	fully connected	layer_size = sequence_length*channels, activation = *ReLu*
		dropout	dropout_probability = 0.3
SBULSTM	Layer 5	Bi-LSTM	units = channels*9
		dropout-wrapper	dropout_probability = 0.3
	Layer 6	Bi-LSTM	units = channels*9
		dropout-wrapper	dropout_probability = 0.3
	Layer 7	Uni-LSTM	units = 150
		dropout-wrapper	dropout_probability = 0.3
Fully Connected and Linear Regression Layers	Layer 8	fully connected	layer_size = 350, activation = *ReLu*
		dropout	dropout_probability = 0.3
	Layer 9	fully connected	layer_size = 150, activation = *ReLu*
		dropout	dropout_probability = 0.3
	Layer 10	Regression	Layer_size = 1

3.3 Two-stage Degradation Model

Through observation of failure feature and literature review [16], there are two obvious partitions in a complete degradation process. The amplitude of vibration signals increases at the first stage and rises rapidly at the second stage until it reaches the failure state in a short time. Thus, the data in the slight and severe degradation stage is fitted into a linear regression model and an exponential model, respectively. In order to detect the point of degradation state transition effectively and automatically, we have predefined

three R^2 values α, β and γ as our thresholds, which correspond to linear regression on window data, exponential regression on window data and linear regression on observed data. The degradation state transition occurs when the R^2 values of the linear regression models are lower than α and γ and that of the exponential model is greater than β.
The exponential random coefficient model can be written as follows:

$$S(t_k) = \Phi e^{\theta(t_k)+\varepsilon(t_k)} + \beta \tag{1}$$

where $S(t_k)$ represents the RMS value at time t_k, θ is a random variable following a prior normal distribution with mean $\mu_0$and variance $\sigma_0{}^2$, ε is an error following a Brownian motion with mean $\mu = 0$and variance σ^2, Φ and $\beta$are the fixed offset determined by the linear regression model since the starting point of the exponential model should be the ending point of the linear model. Then $L(t_k)$can be the logarithm of $S(t_k)$.

An update rule of the prior distribution parameters refers to Bayes' theorem. The detailed computation processes are as follows:

$$\mu_{\theta,t_k} = \frac{\mu_0\sigma^2 + (L(t_k) - \Phi')\sigma_0{}^2}{t_{k\sigma_0{}^2+\sigma^2}} \tag{2}$$

$$\sigma_{\theta,t_k}{}^2 = \frac{\sigma^2\sigma_0{}^2}{t_{k\sigma_0{}^2+\sigma^2}} \tag{3}$$

Prior to determine the distribution of the failure time, we have predefined the failure threshold $\delta = 0.30$. Once degradation value RMS reaches the failure threshold, we assume the failure occurs. Then, the mean and variance of $L(t_k + t)$can be determined as:

$$\underline{\mu}(t_k + t) = \mu_{\theta,t_k}t + L(t_k) \tag{4}$$

$$\underline{\sigma}^2(t_k + t) = \sigma_{\theta,t_k}{}^2 + \sigma^2 t \tag{5}$$

Eventually, the conditional cumulative distribution function of T, the remaining life, can be computed as follows:

$$\begin{aligned} P(T \leq t|L(t_1),\dots,L(t_k)) &= P(L(t_k + t) \geq \delta|L(t_1),\dots,L(t_k)) \\ &= 1 - P(L(t_k + t) \leq \delta|L(t_1),\dots,L(t_k)) \\ &= 1 - P(Z \leq \tfrac{\delta-\underline{\mu}(t_k+t)}{\underline{\sigma}(t_k+t)}|L(t_1),\dots,L(t_k)) \\ &= \Phi(\tfrac{\underline{\mu}(t_k+t)-\delta}{\underline{\sigma}(t_k+t)}) \end{aligned} \tag{6}$$

To correct the posterior mean for θ at time t_k, it can be determined as follows:

$$\Delta\theta(t_k) = K_c e(t_k) \tag{7}$$

where K_c represents correction gain and $e(t_k)$is an error term for the correction. With $\Delta\theta(t_k)$, we are able to update the mean of θ in the prior distribution.

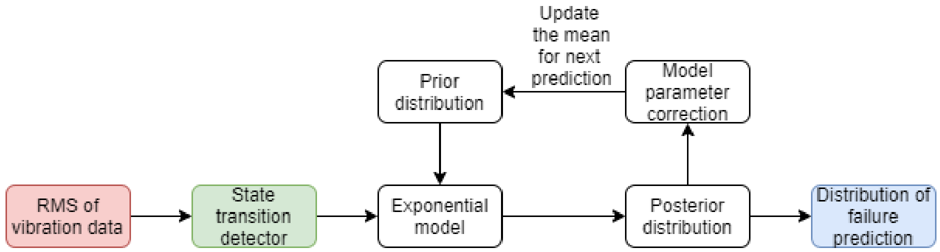


Figure 2. Framework of the proposed two-stage degradation approach.

3.4 Metrics for Determining RUL

Based on literature review, there are two common ways to predict RUL. One is to model RUL as linear relationship with respect to effective machine rotary running time, $RUL = RUL_{max} - t$, and then uses multi-features as input to predict RUL directly [15]. The other way is to predict selected features and set threshold bars manually. Once the feature predictions reach thresholds, the cross point is the predicted RUL. Both metrics were used separately in the experiment.

4. Experimental Results

4.1 Data Extraction & Feature Selection

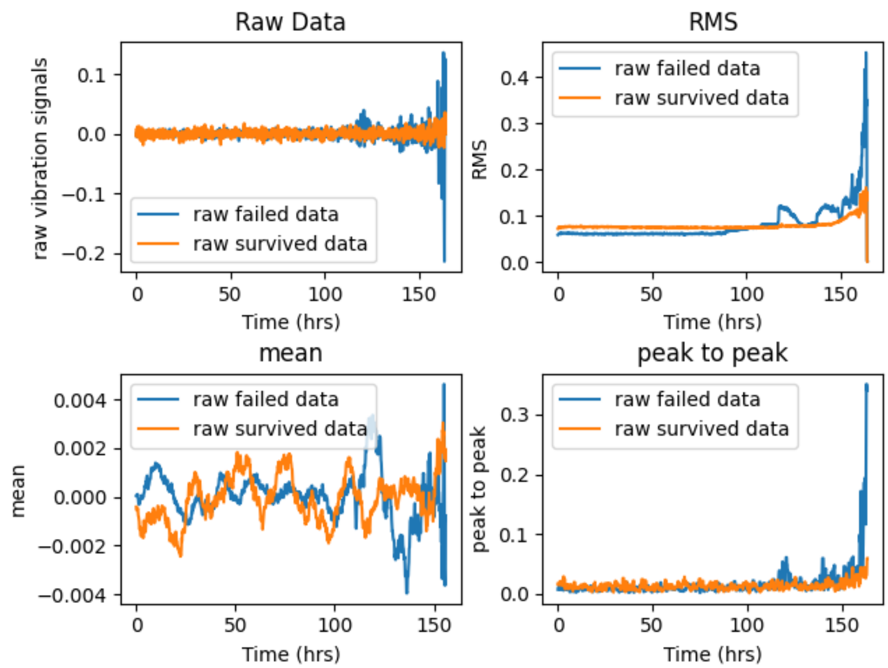


Figure 3. Raw data of bearing test-to-failure experiment.

Based on raw vibration signals, the potential features are root of mean square (RMS), mean and peak to peak. The failed bearings vibrate heavily and had higher RMS and peak to peak values when approaching the end of the experiment compared to survived bearings. Therefore, both RMS and peak to peak could be good health indicators for both RUL and anomaly detection. The mean feature is not considered because it does not show a strong correlation with bearing's degradation.

4.2 Experimental Results

4.2.1 LSTM

With three sets of RMS features extracted from raw dataset, a LSTM model was trained on two of them and tested on the third set. Test result of three broken bearings is shown in Fig Y. From the plot we learned that bearing 3 has a smooth transition to the final rise of RMS while bearing 2 and 3 have fluctuations during the start and before the final rise respectively. In this case, the RMS failure threshold is set to be 0.3. Only the first case reached the threshold. Thus, the LSTM model failed to capture the rapid rising trend of RMS in the severe degradation stage. Similar results obtained by using peak to peak as features.

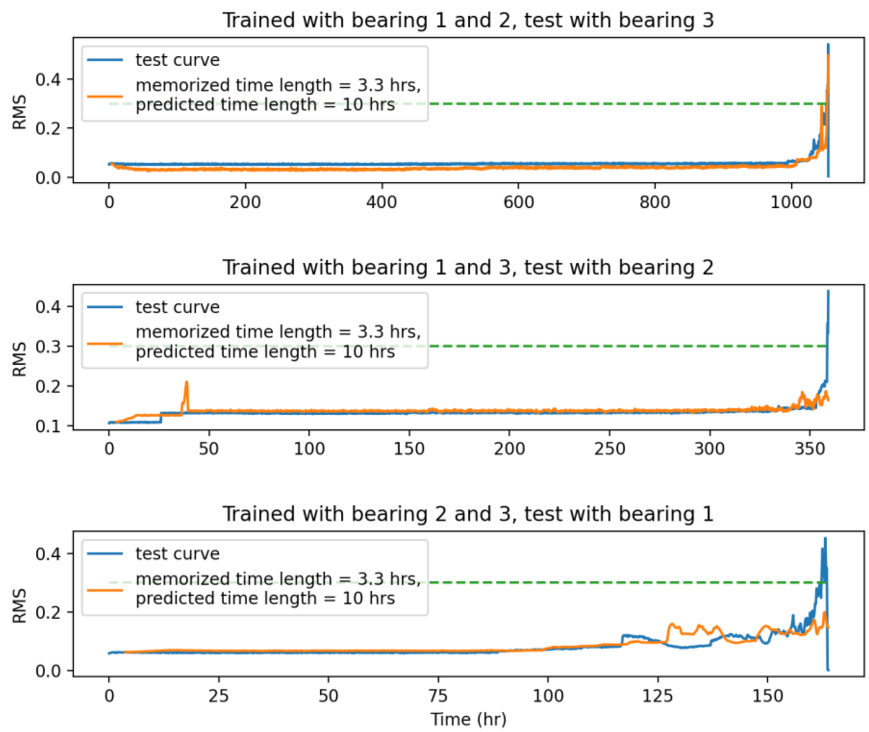


Figure 4. Plots of predictions on all three sets of failed bearing data.

4.2.2 CNN + SBULSTM

As suggested by Fig 4, the pure LSTM model fails to predict the rising trend of RMS and peak to peak values when working with noisy input. The RMS of bearings in Fig 4 has multiple small bumps and pulses before the final rise, which mislead our LSTM model. After tuning the hyperparameter, the result does not meet our expectations. Therefore, another innovative framework CNN + SBULSTM was applied. Fig 5 shows the model's predictions on bearing 1's RMS and bearing 3's peak to peak values. When working with RMS, our model predicted the rising trend however the error of the RUL computed is still very large. Peak to peak dataset has more noise than RMS, even with SBULSTM, the model failed to make sensible results. As shown in the Fig.5, the prediction of peak to peak reaches the failure threshold much earlier than actual failure time.

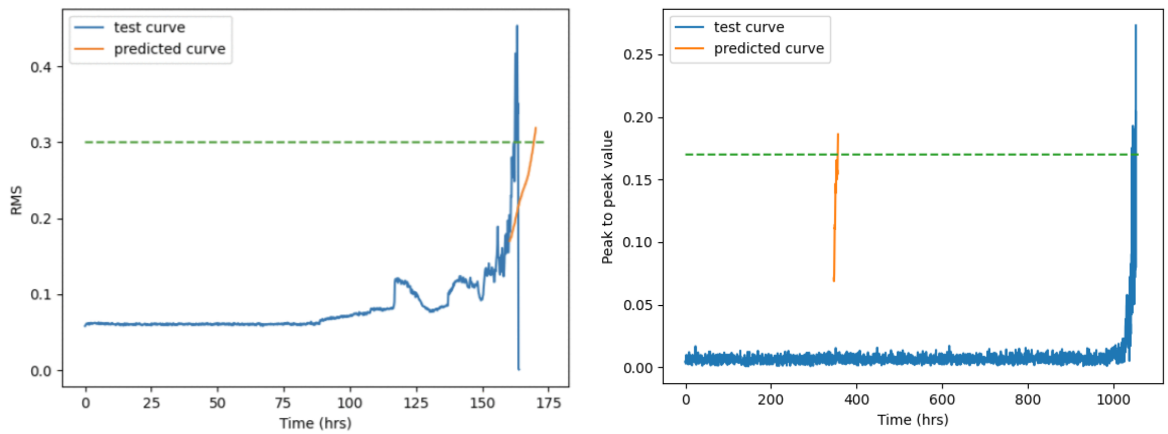


Figure 5. 1D CNN + SBULSTM prediction on RMS and peak to peak prediction of test 2 bearing 1 case.

Instead of using failure threshold to find RUL, directly predicting RUL with RMS and peak to peak as inputs to the machine learning framework was also tried. However, due to abnormality of both features such as large noise and weak correlation between features and RUL, the output cannot capture the change of RUL.

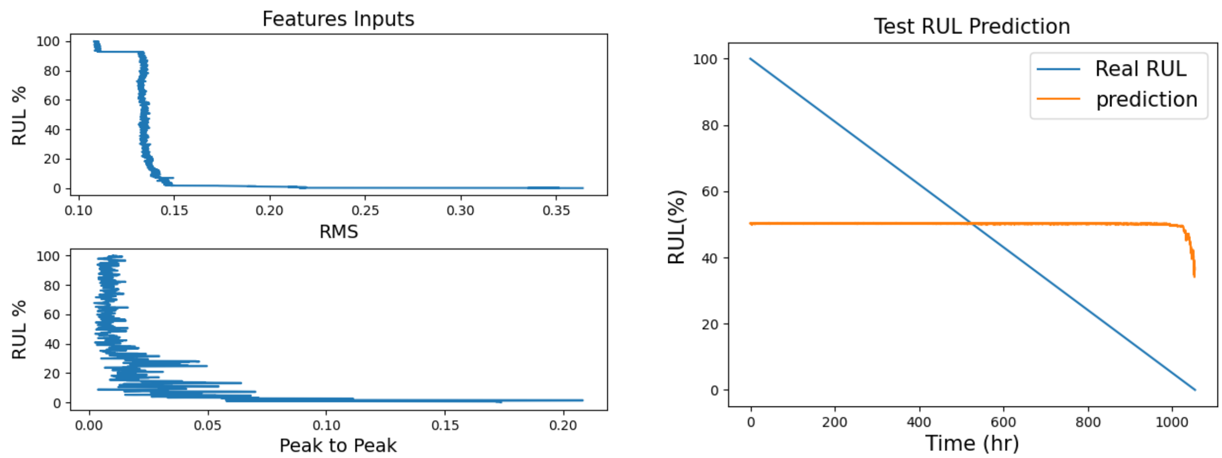


Figure 6. 1D CNN + SBULSTM RUL prediction on test 2 bearing 1 case.

Since CNN and SBULSTM are data-hungry machine learning models, the main reason that our model didn't perform well is due to the lack of large dataset with less noise. We believe that with more training data, this model could behave much better.

4.2.3 Two-stage Degradation Prediction Model

Based on our proposed two-stage degradation model, three failure cases are predicted. After tuning each value of threshold α, β and γ, slight and severe degradation stages of RMS data are fitted by linear regression and random exponential random coefficient model respectively. For each instance, fitting models of real time RUL prediction is updated by incorporating newly detected data points and it will predict future RMS in the next 3.5 hours. The Figure 7 shows the results after RUL reaching 3 hours for both cases and corresponding errors are shown in Table 2. With an average prediction error of 0.45 hour, the proposed model is believed to give acceptable forecast results for bearing in rotary machines with little wastes in industry. Peak to peak features were not used because it is too noisy, and the two stages cannot be detected accurately.

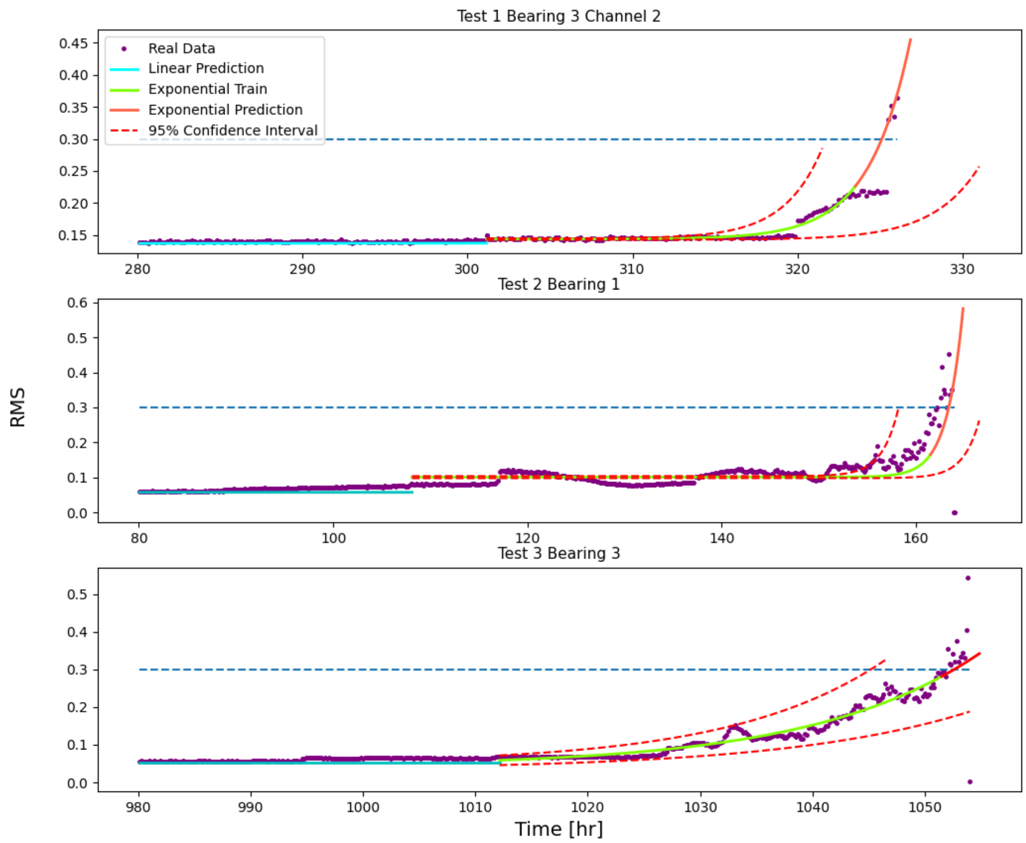


Figure 7. Two-stage degradation prediction model for all three failure cases.

Table 2. Two-stage degradation prediction accuracy

Rolling Bearing Dataset	Real RUL [hr]	Predicted RUL [hr]	Predict Error [hr]
Test 1 Bearing 3	3.00	2.70	0.30
Test 2 Bearing 1	3.00	2.45	0.55
Test 3 Bearing 3	3.00	2.55	0.45

6. Conclusion:

In this project, a novel intelligent RUL prediction method - two-stage degradation model is proposed and implemented on a natural run-to-failure bearing datasets. Accurate RUL prediction is obtained by our proposed two-stage degradation model within an acceptable time interval. It offers better description of bearing failure features and exhibits superior performance in RUL prediction compared to RNN+LSTM and CNN+LSTM frameworks. This method can be applied to general cases and identify the bearing failure stages in industrial applications.

Reference

[1] Zhang, Shen, et al. "Deep learning algorithms for bearing fault Diagnostics—A comprehensive review." *IEEE Access* 8 (2020): 29857-29881.
[2] Blodt, Martin, et al. "Models for bearing damage detection in induction motors using stator current monitoring." *IEEE transactions on industrial electronics* 55.4 (2008): 1813-1822.
[3] Lopez-Perez, David, and Jose Antonino-Daviu. "Application of infrared thermography to failure detection in industrial induction motors: case stories." *IEEE Transactions on Industry Applications* 53.3 (2017): 1901-1908.
[4] Esfahani, Ehsan Tarkesh, Shaocheng Wang, and V. Sundararajan. "Multisensor wireless system for eccentricity and bearing fault detection in induction motors." *IEEE/ASME Transactions on Mechatronics* 19.3 (2013): 818-826.
[5] Harmouche, Jinane, Claude Delpha, and Demba Diallo. "Improved fault diagnosis of ball bearings based on the global spectrum of vibration signals." *IEEE Transactions on Energy Conversion* 30.1 (2014): 376-383.
[6] M.-Y. Chow, P. Mangum and S. Yee, "A neural network approach to real-time condition monitoring of induction motors", IEEE Trans. Ind. Electron., vol. 38, no. 6, pp. 448-453, Dec. 1991.
[7] A. Malhi and R. Gao, "PCA-based feature selection scheme for machine defect classification", IEEE Trans. Instrum. Meas., vol. 53, no. 6, pp. 1517-1525, Dec. 2004.
[8] D. He, R. Li, J. Zhu and M. Zade, "Data mining based full ceramic bearing fault diagnostic system using AE sensors", IEEE Trans. Neural Netw., vol. 22, no. 12, pp. 2022-2031, Dec. 2011.
[9] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position", Biol. Cybern., vol. 36, no. 4, pp. 193-202, Apr. 1980.
[10] J. Schmidhuber, "Deep learning in neural networks: An overview", Neural Netw., vol. 61, pp. 85-117, Jan. 2015.
[11] Y. Bengio, "Learning deep architectures for AI", FNT Mach. Learn., vol. 2, no. 1, pp. 1-127, 2009.
[12] W. Abed, S. Sharma, R. Sutton and A. Motwani, "A robust bearing fault detection and diagnosis technique for brushless DC motors under non-stationary operating conditions", J. Control Autom. Electr. Syst., vol. 26, no. 3, pp. 241-254, Jun. 2015.
[13] W. Mao, et al. "Predicting remaining useful life of rolling bearings based on deep feature representation and long short-term memory neural network." Advances in Mechanical Engineering 10.12 (2018): 1687814018817184.
[14] S. Selvin, R. Vinayakumar, E. A. Gopalakrishnan, V. K. Menon and K. P. Soman, "Stock price prediction using LSTM, RNN and CNN-sliding window model," 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Udupi, India, 2017, pp. 1643-1647, doi: 10.1109/ICACCI.2017.8126078.
[15] S. Zhang, S. Zhang, B. Wang and T. G. Habetler, "Deep Learning Algorithms for Bearing Fault Diagnostics—A Comprehensive Review," in IEEE Access, vol. 8, pp. 29857-29881, 2020, doi: 10.1109/ACCESS.2020.2972859.
[16] Z. Pan, Z. Meng, Z. Chen, W. Gao and Y. Shi, "A two-stage method based on extreme learning machine for predicting the remaining useful life of rolling-element bearings", Mechanical Systems and Signal Processing, Vol. 144, Oct. 2020
[17] J. Moyne and J. Iskandar, "Big Data Analytics for Smart Manufacturing: Case Studies in Semiconductor Manufacturing", MDPI, 2017

[18] Y. Chen, T. Shou, W. Li, C. He, B. Yang, "Predicting Rolling Bearing Remaining Useful of Life based on Data-driven Methods", UMich EECS 545 Progress Report, 2021
[19] Q. An, Z. Tao, X. Xu, M. Mansori, M. Chen. "A data-driven model for milling tools remaining useful life prediction with convolutional and stacked LSTM network". Measurement, Volume 154, 2020.

Author Contribution

"Wenxi and Chuan collected data and worked on pre-processing. Yizhou and Tianyu developed classification pipelines and performed hyperparameter search. Bingqian helped with data collection, experiments, and error diagnostics. All co-authors were involved in writing this report. All co-authors equally contributed to this project."

Appendix

Dataset link: https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/
Github link: https://github.com/tianyushou/bearingRUL