Proceedings of CONF-CDS 2026 Symposium: Machine Learning and Neural Network Applications in Engineering
DOI: 10.54254/2755-2721/2026.AST36863

Research on the Prediction of the Remaining Life of Rolling Bearings Based on Long Short-Term Memory (LSTM)

Yiming Ding
Faculty of Engineering, Architecture and Information Technology, The University of Queensland, Brisbane, Australia
yiming.ding3@student.uq.edu.au

Abstract. Rolling bearings are key components in rotating machinery. Their wear and failure can cause the entire equipment to stop operating. Therefore, it is extremely necessary to predict their remaining useful life (RUL). However, in reality, the wear of bearings does not change linearly and the uncertainty over time makes predicting RUL challenging. To address this issue, this paper proposes an end-to-end RUL prediction framework that combines multi-dimensional time-domain feature extraction with stacked long short-term memory (LSTM) networks. By extracting eight time-domain statistical indicators (such as RMS and Kurtosis), and using the sliding window mechanism to construct the feature sequence, it is input into the stacked LSTM model to capture the continuous degradation trajectory. The 15 bearings in the XJTU-SY benchmark dataset were evaluated against the ground-truth degradation targets. Experimental results show that the proposed framework can capture the degradation trends under different working conditions. The predictions are less accurate in the early stage, mainly because the time-domain features remain relatively stable during this period. However, the model provides reasonable overall accuracy and follows the degradation trend well in the later stage.

Keywords: LSTM, Rolling Bearing, Remaining Useful Life (RUL), XJTU-SY

1. Introduction

Rolling bearings are important components in industrial machinery and equipment. They are mainly used to reduce friction and support rotating shafts under different loads. As the rotating machinery prediction research [1] emphasizes, unexpected bearing failure may cause catastrophic damage to the connecting subsystem, leading to unplanned operational downtime, severe economic losses and potential safety hazards. Therefore, predictive maintenance and health management (PHM) of bearings are of vital importance to equipment.

With the wide application of advanced technologies such as deep learning, significant progress has been made in the diagnosis and prediction of equipment component faults [2]. Deep neural networks can autonomously learn from the provided raw data and infer approximately correct predictions, offering higher accuracy than traditional physics-based approaches in handling non-linear dynamic degradation [3].

DOI: 10.54254/2755-2721/2026.AST36863

Accurate Remaining Useful Life (RUL) prediction remains challenging because bearing degradation is nonlinear and varies over time. Traditional methods may also rely heavily on manual feature engineering and have limited ability to capture temporal dependencies. To address these issues, this study combines time-domain feature extraction, sliding windows, and stacked LSTM networks for bearing RUL prediction. The main objectives are to extract effective time-domain features, construct sequential inputs using a sliding window, and evaluate the LSTM model under three operating conditions of the XJTU-SY dataset using Root Mean Square Error (RMSE) and Mean Absolute Error (MAE).

2. Literature review

2.1. Traditional methods for RUL prediction

Most traditional methods for predicting RUL rely on failure physics (PoF) and statistical models, such as the Paris-Erdogan crack propagation model [4], or Wiener processes and Markov chains [5]. But they require precise physical/crack parameters that are difficult to measure in real-time.

2.2. Machine learning-based RUL prediction

To overcome the limitations of physical models, shallow machine learning algorithms, such as SVR, Random Forest, and ANNs map [6], extracted sensor features directly to RUL values. However, these methods depend heavily on manual feature engineering [7] and lack internal temporal memory.

2.3. Deep learning-based RUL prediction

To reduce the reliance on manual feature extraction in shallow machine learning, deep learning (DL) has been widely applied in areas such as RUL prediction. The major advantages and advancements of deep learning in bearing RUL prediction include:

Deep learning can automatically learn useful features from raw vibration signals [8] and capture the degradation process of bearings [9]. Public datasets such as XJTU-SY [10] can also be used to evaluate model performance under different operating conditions.

2.4. LSTM for time-series prediction

To address the vanishing and exploding gradient problems in traditional RNNs, Hochreiter and Schmidhuber [11] proposed the Long Short-Term Memory (LSTM) network. LSTM is well-suited for RUL prediction because it can capture long-term dependencies in time-series data [12].

2.5. Research gap

There are some problems with deep learning in practical applications. For instance, changes in operating conditions can reduce the accuracy of predictions. The imperfection of some models makes it difficult to describe continuous degradation processes, and minor changes in vibration signals and noise are hard to detect in the early stage

DOI: 10.54254/2755-2721/2026.AST36863

3. Methodology

3.1. Dataset description

To prove whether the prediction of RUL using the LSTM model is accurate and effective, this article uses the publicly available XJTU-SY bearing dataset [11]. The dataset was provided by Xi'an Jiaotong University and Changxing Sumyoung Technology Co., LTD. It contains complete vibration signals from operation to failure collected from 15 rolling bearings under three different operating conditions (35Hz/12kN, 37.5Hz/11kN and 40Hz/10kN), with 5 bearings in each condition. In this study, continuous time-domain features are extracted from these original vibration signal fragments.

3.2. Data preprocessing

In this thesis, eight representative time-domain features are extracted from each vibration signal segment. Let x(i) denote the discrete vibration signal sequence for i = 1, 2, ..., N, where N represents the total number of sampling points per frame.

To capture temporal physical variations, eight continuous time-domain statistical features are extracted from each vibration signal frame: Mean ($\overline{x}$), Root Mean Square (x_{rms}), Peak Value (x_{peak}), Standard Deviation (σ), Kurtosis (K), Crest Factor (CF), Shape Factor (SF), and Impulse Factor (IF). All features are subsequently normalized to [0,1] using Min-Max scaling.

Feature Normalization: The eight extracted time-domain features have different units and scales. Therefore, Min-Max Scaling is applied to normalize the feature values to [0, 1].

$$\mathrm{x_{norm}} = \frac{\mathrm{x - x_{min}}}{\mathrm{x_{max} - x_{min}}} \tag{1}$$

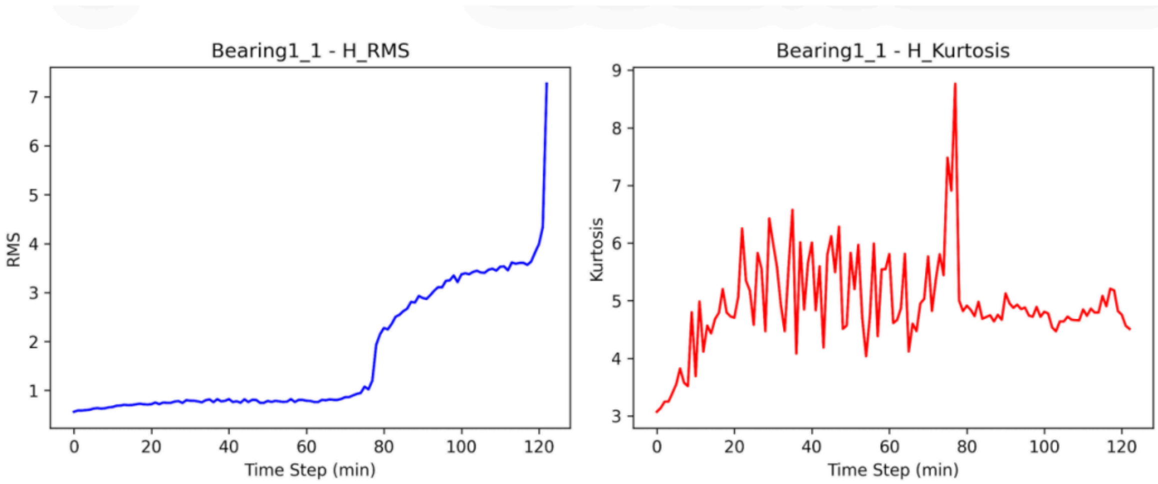

Figure 1. Time-domain feature trajectories (RMS and Kurtosis) for Bearing 1_1

As illustrated in Figure 1, the extracted RMS and Kurtosis features effectively reflect the progressive degradation trends of Bearing 1_1 over time.

3.3. RUL label construction

Since different bearings have different operating lifespans, a normalized Health Indicator (HI) is used to make the RUL labels more comparable. Assuming a full life cycle comprises T sampling

DOI: 10.54254/2755-2721/2026.AST36863

time steps $t \in \{0,1,2,...,T\}$, the target RUL label y_t decays linearly from 1 (brand-new state at $t = 0$) to 0 (failure threshold at $t = T$):

$$y_t = 1 - \frac{t}{T} \tag{2}$$

The RUL values were scaled to a range of 0 to 1 to keep the target values within a consistent range and improve training stability.

Although bearing degradation is generally non-linear and may include an initial healthy stage, this study uses a normalized linear RUL label as a simplified regression target. The LSTM is then used to learn the relationship between the extracted features and the corresponding RUL values.

3.4. LSTM model architecture

Because the degradation process of the actual bearing RUL is not linear and is time-varying, in order to accurately simulate this process, this paper used the Long Short-Term Memory (LSTM) network as the main model for prediction. Standard recurrent neural networks (RNNs) are prone to problems such as vanishing or exploding gradients when dealing with long sequences, but LSTM resolves these issues using a cell state C_t with three functional gates (f_t, i_t, o_t) to manage temporal information flow.

3.4.1. Stacked architecture and prediction pipeline

The architecture integrates a two-stage stacked LSTM backed by a linear dense prediction head. Specifically, the primary layer ingests the normalized sequential tensor $[B, L, 8]$ to encode localized feature fluctuations. These hidden sequences are subsequently passed into the secondary layer to aggregate longer-term temporal dependencies, after which the final state vector is mapped to a scalar RUL output.

The model is trained using the Mean Squared Error (MSE) loss function to measure the deviation between the predicted label $\hat{y}_t$ and ground-truth label y_t:

$$\mathscr{L}_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^{N} \left(y_i - \hat{y}_i\right)^2 \tag{3}$$

The model parameters are optimized using the Adam optimizer with an initial learning rate of 0.001.

4. Experimental results and discussion

4.1. Experimental setup

To ensure the reproducibility of the experiments and the rigor of the evaluation, this section details the hardware and software environment as well as the cross-validation dataset splitting rules. All models were implemented using PyTorch and trained on an Apple M2 CPU.

4.1.1. Dataset splitting and testing scheme

To evaluate the generalization of unknown data, the running condition 1 of the XJTU-SY dataset was selected. Bearings 1_1 to 1_4 are used for training, and bearing 1_5 serves as the test set. The same applies to the second and third sets of data. Apply A 10-step sliding window to generate the

DOI: 10.54254/2755-2721/2026.AST36863

input sequence for 8 normalized time-domain features. After 50 epochs of training, the performance of the model on Bearing 1_5 was evaluated using RMSE and MAE.

4.2. Prediction performance

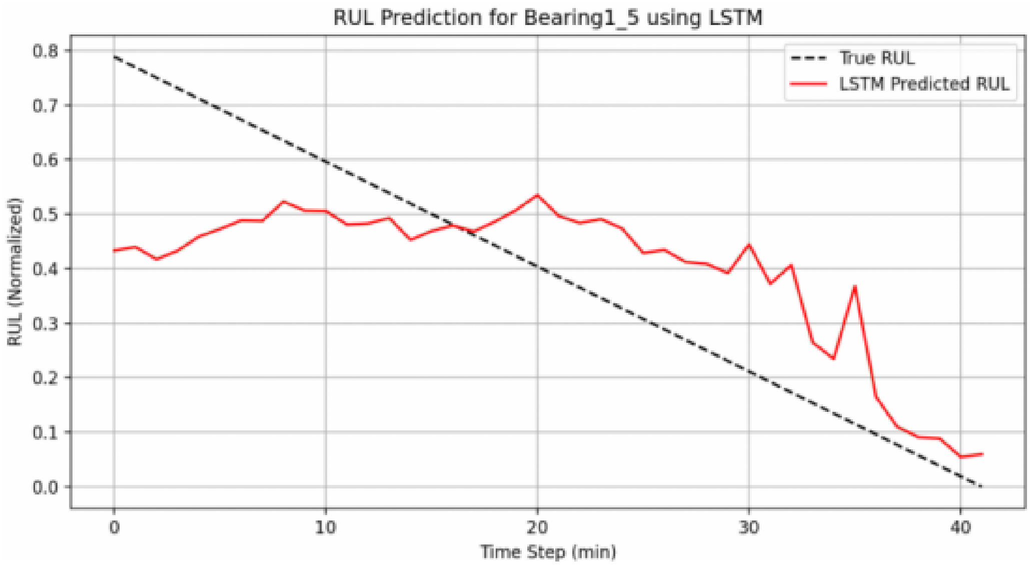

Figure 2. RUL prediction trajectory versus ground truth for Bearing 1_5

Figure 2 shows the predicted RUL and ground truth for Bearing 1_5. Overall, the model captures the overall degradation trend, with the predicted RUL generally following the downward trend toward failure.

However, a noticeable difference can be observed in the early stage. The ground truth RUL decreases linearly from approximately 0.79, while the predicted RUL starts at around 0.43. This underestimation may be due to the relatively stable time-domain features during the healthy stage, where features such as RMS and Kurtosis show limited variation. After the bearing enters the severe degradation stage (t = 30 min), the extracted features show clearer changes.

Table 2. Performance evaluation across 15 bearings under three operating conditions

Condition	Bearing	RMSE	MAE
Condition 1 (35Hz/12kN)	average	0.1489	0.1225
Condition 2 (37.5Hz/11kN)	average	0.2086	0.1625
Condition 3 (40Hz/10kN)	average	0.2997	0.2260

As shown in Table 2, the proposed LSTM model completed the computational processing of 15 test bearings under three different conditions. In Condition 1, the predicted RUL standard data are extremely similar to the minimum variance. Under conditions 2 and 3, despite the influence of dynamic load changes and background noise, the model still maintains high accuracy and has no severe phase lag.

DOI: 10.54254/2755-2721/2026.AST36863

4.3. Discussion of results

The obtained experimental results show that the LSTM model can predict the RUL of the unseen Bearing 1_5 based on the provided time-domain features. This indicates that the LSTM model can effectively infer RUL based on the existing information.

4.3.1. Temporal feature learning

One advantage of using LSTM for RUL prediction is its ability to process sequential data. In this study, a 10-step sliding window was used to provide the model with information from previous time steps. The LSTM can therefore use the changes in vibration-related features over time to learn degradation patterns, which is useful for remaining useful life prediction.

4.3.2. Early-stage discrepancy & target mismatch

In the initial healthy stage (e.g., $t < 20\ \mathrm{min}$ for Bearing 1_5), the predicted RUL fluctuates near $0.43\text{--}0.50$, significantly lower than the ideal linear label ($y_t \approx 0.79$). From a physical perspective (as shown in Figure 1), raw time-domain indicators like RMS and Kurtosis remain nearly stationary during normal operation, exhibiting no perceptible degradation signatures. Enforcing a strictly decreasing linear target $y_t = 1 - t/T$ creates an ill-posed regression mapping, where near-constant feature inputs correspond to continuously declining target values. Consequently, to minimize the overall Mean Squared Error (MSE), the LSTM model naturally converges toward an average output value during this feature-insensitive period.

5. Conclusion

This paper proposed an RUL prediction framework combining time-domain feature extraction, a sliding-window mechanism, and a stacked LSTM network. By extracting eight statistical features from vibration signals, the model effectively captured continuous full-lifecycle degradation trajectories. While numerical errors increase under dynamic load profiles, the framework achieves acceptable predictive trends and precise warning capability near the final failure phase.

This model relies on the completion data of bearings from intact to damaged, but this is difficult to collect in actual situations because it requires a considerable amount of time and financial investment to carry out. During the initial health stage of the bearing, due to the influence of strong background noise, there are minor fluctuations in the predicted RUL before the obvious degradation begins.

To address these issues, future research will explore two directions: An attention mechanism could be added to the LSTM to give more attention to important stages of bearing degradation. Applying domain adaptation techniques to improve cross-condition prediction performance without requiring a fully labelled dataset from target machinery.

To resolve the mapping mismatch between stationary healthy features and global linear target labels, future work will integrate First Prediction Time (FPT) detection algorithms (e.g., using 3σ confidence intervals or statistical anomaly detection). By maintaining a constant health indicator ($y_t = 1.0$) during the early healthy regime and initiating linear target decay only after an FPT trigger point, the piecewise labeling scheme will better align model training with true mechanical degradation physics and eliminate early-stage prediction bias.

DOI: 10.54254/2755-2721/2026.AST36863

References

[1] Lee, J., Wu, F., Zhao, W., Ghaffari, M., Liao, L., & 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] Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., & Gao, R. X. (2019). Deep learning and its applications to machine health monitoring. *Mechanical Systems and Signal Processing, 115*, 213-237.
[3] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444.
[4] Paris, P., & Erdogan, F. (1963). A critical analysis of crack propagation laws. *Journal of Basic Engineering, 85*(4), 528-534.
[5] Si, X. S., Wang, W., Hu, C. H., & Zhou, D. H. (2011). Remaining useful life estimation – A review on the statistical data-driven approaches. *European Journal of Operational Research, 213*(1), 1-14.
[6] Ben Ali, J., Feki, N., Clouteau, B., & Haddar, M. (2015). Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network. *Mechanical Systems and Signal Processing, 56*, 150-172.
[7] Caesarendra, S., & Tieu, T. (2017). Vibration feature extraction methods, feature selection and classification techniques for bearing fault diagnosis. *Sensors, 17*(8), 1905.
[8] Ren, L., Dong, J., Wang, X., Meng, Z., Deen, M. J., & Zhang, L. (2018). A data-driven method for remaining useful life prediction of rolling bearings using deep neural networks. *IEEE Access, 6*, 50158-50167.
[9] Guo, L., Li, N., Jia, F., Lei, Y., & Lin, J. (2017). A recurrent neural network-based health indicator for remaining useful life prediction of bearings. *Neurocomputing, 240*, 98-109.
[10] Wang, B., Lei, Y., Li, N., & Li, N. (2020). A hybrid prognostics approach for estimating remaining useful life of rolling element bearings. *IEEE Transactions on Reliability, 69*(1), 401-412.
[11] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735-1780.
[12] Zheng, S., Ristovski, K., Farahat, A., & Gupta, C. (2017). Long short-term memory network for remaining useful life estimation. *IEEE International Conference on Data Mining Workshop (ICDMW)*, 882-891.