



Machine Learning Framework For Predicting Machine State And Remaining Useful Life

¹Syed Azadar Hussain Zaidi, ²Dr. Pankaj Prajapati

¹MTech Research Scholar, ²Professor,

¹Department of Robotics

¹Ambalika Institute of Technology, Lucknow

Abstract: This paper presents a novel machine learning framework for predicting the health state and remaining useful life (RUL) of industrial machinery and equipment. The proposed approach addresses key challenges in data-driven prognostics, including feature engineering, temporal modelling, and uncertainty quantification. Case studies on wind turbine main bearings and milling machine tool wear further validate the framework's practical applicability and generalizability. This research advances the field of machine prognostics by providing a flexible, accurate, and interpretable framework for predicting machinery health and RUL. The proposed approach enables more effective predictive maintenance strategies, potentially reducing downtime and maintenance costs in various industrial applications.

Index Terms - Predictive Maintenance, RUL, Machine Prognostics, Machine Learning

1. INTRODUCTION

For Machine prognostics and health management (PHM) is a critical area that enables predictive maintenance of industrial assets. By monitoring the health state of machines and predicting their remaining useful life (RUL), unplanned downtime can be minimized while optimizing maintenance costs. The increasing availability of sensor data from machines, combined with advancements in machine learning, has opened up new opportunities for data-driven approaches to machine prognostics.

However, building effective machine learning models for RUL prediction poses several challenges. First, the degradation patterns of machines can be complex and non-linear, requiring models that can learn sophisticated relationships from data. Second, sensor data is often high-dimensional and noisy, necessitating feature engineering or automated feature learning from raw signals. Third, many real-world scenarios involve limited failure data, as running machines to failure can be prohibitively expensive. This motivates techniques that can learn from limited labeled data, such as transfer learning. Finally, for practical deployment, the models need to be computationally efficient and provide well-calibrated uncertainty estimates.

This dissertation aims to address these challenges by developing a novel machine learning framework for predicting machine health state and RUL. The proposed framework leverages unsupervised representation learning to extract degradation-related features from raw sensor data. These learned representations are then used as inputs to a dynamic neural network architecture that can model temporal dependencies and provide uncertainty-aware predictions. The framework is validated on two prognostics case studies - bearing RUL prediction and aircraft engine RUL prediction. Experiments demonstrate improved accuracy and computational efficiency compared to state-of-the-art benchmarks.

2. LITERATURE REVIEW

P. Nectoux et al. has introduced PRONOSTIA, a platform built for testing and validating different methods for assessing the health of bearings, which are essential parts in rotating machines. Since bearing failures can greatly reduce machine safety and availability, this platform aims to provide data on how bearings degrade over time. PRONOSTIA uses sensors to monitor operating conditions like rotating speed and force, and it collects temperature and vibration data that help track the bearing's health in real time. This data can then be analyzed to extract important features related to bearing degradation. Additionally, in the IEEE PHM 2012 Prognostic Challenge, competitors used PRONOSTIA's data to try out their prognostic methods for predicting how much life is left in bearings [2]. A. Saxena et al. has explored methods for modeling damage progression in aircraft gas turbine engine modules. By using a thermodynamic simulation model, sensor response surfaces are generated based on variations in flow and efficiency across different engine modules. Each dataset imposes an exponential change rate on flow and efficiency, beginning from a randomly selected starting point of deterioration. This approach aims to simulate how damage might develop within the engine, providing insights into predictive maintenance strategies[3]. Tool wear significantly impacts workpiece quality in the machining industry, making accurate wear prediction essential for cost-effective operations. T. Benkedjouh et al. suggests that parameters like cutting force, vibration, and acoustic emission (AE) play critical roles in tool wear. By effectively predicting wear, tools can be replaced as needed, helping to minimize costs and improve product quality [4]. J. Lee et al offers a detailed review of the field of Prognostics and Health Management (PHM), and introduces a structured design approach called the 5S methodology. This method helps transform raw data into useful prognostic information by identifying key components and providing tools to choose the most suitable algorithms for different applications. The approach aims to streamline PHM design for more effective implementation in diverse settings. [5]

3. MATERIALS AND EXPERIMENT METHODS

3.1 Unsupervised Feature Learning with RAP Clustering

Feature extraction from raw sensor data is considered to be very vital in data-driven machine prognostics. Feature construction has been earlier done traditionally by hand-picking features from domain knowledge and applying statistical or signal processing algorithms to construct health indicators. However, these hand-crafted features might not be suitable for different machines or even different operating conditions. Consequently, manual feature extraction remains a trade off, as it is timely, and more so requires extra attention to get the best for every new data set. [17]

3.2 Reduced Affinity Propagation (RAP) Clustering

This section describes the proposed RAP clustering algorithm in detail. First, the original AP algorithm is briefly reviewed, followed by the modifications introduced in RAP to handle time series data.[7]

3.2.1 Affinity Propagation (AP) Clustering [18]

AP then initializes two matrices, the responsibility matrix $R \in \mathbb{R}^{N \times N}$ and the availability matrix $A \in \mathbb{R}^{N \times N}$, to zero. The responsibility r_{ij} represents the accumulated evidence for data point x_j to be chosen as the exemplar for data point x_i , while the availability a_{ij} represents the accumulated evidence for data point x_i to choose data point x_j as its exemplar. AP iteratively updates the responsibility and availability matrices using the following update rules:

$$r_{ij} \leftarrow s_{ij} - \max_{k \neq j} (a_{ik} + s_{ik})$$

$$a_{ij} \leftarrow \min \left(0, r_{jj} + \sum_{k \notin \{i, j\}} \max(0, r_{kj}) \right)$$

The responsibility update rule computes the evidence for data point x_j to be the exemplar for data point x_i by subtracting the maximum evidence for other potential exemplars. The availability update rule computes the evidence for data point x_i to choose data point x_j as its exemplar by considering the self-responsibility r_{jj} and the positive responsibilities of other data points towards x_j . AP iterates between the responsibility and availability updates until convergence or a maximum number of iterations is reached. The exemplars are then identified as the data points with positive self-responsibility, i.e., the data points with $r_{ii} > 0$.

AP has several advantages over traditional clustering algorithms like k-means:

1. The number of clusters is automatically determined based on the data and the preferences, without requiring it to be specified in advance.
2. By considering all data points as potential exemplars, AP can find clusters of varying sizes and shapes that capture the intrinsic structure of the data.
3. The exemplars are actual data points, providing a meaningful representation of each cluster.
4. AP can be applied to problems where the similarities are not symmetric and do not satisfy the triangle inequality.

While AP are not without limits:

1. AP has a computational complexity of $O(N^2)$ per iteration (where N is the number of data points), making it infeasible for very large datasets.
2. Results will be sensitive to the preferences chosen, which underlie any coarse-graining and also require some additional knowledge of the data. The median of similarities is used in most cases, but some domain knowledge can be required to get proper thresholds.
3. AP does not output a hierarchical structure of clusters, which is required in some applications.

3.2.2 Reduced Affinity Propagation (RAP) [19]

Although AP is successful in finding clusters with very good quality, computational cost of its iteration $O(N^2)$ makes it impracticla for some large datasets. In order to achieve this goal, we introduce Reduced Affinity Propagation (RAP) which is a new variant of additive propagation (AP), and it dramatically decreases the computational cost while not causing much damage on clustering performance. This is done to fix that we identified the restraint.

RAP is based on the idea that each data point should not have a large number of potential exemplars to consider. This is done simply by using a nearest neighbor graph. More specifically, we first find k nearest neighbours of each data point x_i according to similarity matrix S . We then only allow these points to be candidate exemplars for x_i . This process has been repeated for each data point. Because of this, the amount of messages that need to be exchanged during each iteration is reduced from N^2 to Nk , where k is less than infinity.

The RAP algorithm proceeds as follows:

1. Compute the similarity matrix S between all pairs of data points.
2. For each data point x_i :
 - Find the k nearest neighbors of x_i based on S .
 - Initialize the responsibility $r(i,j)$ to $s(i,j) - \max_{\{k \neq j\}} s(i,k)$ for j in the k nearest neighbors of i , and to $-\infty$ otherwise.
 - Initialize the availability $a(i,j)$ to 0.
3. Repeat until convergence:
 - Update the responsibilities $r(i,j)$ for each x_i and its k nearest neighbors x_j using the standard AP update rule.
 - Update the availabilities $a(i,j)$ for each x_i and its k nearest neighbors x_j using the standard AP update rule.
4. Assign each data point to the exemplar that maximizes the sum of responsibility and availability, considering only the k nearest neighbor candidates.

The combination of RAP-based exemplar learning and ESN-based RUL prediction enables the proposed approach to effectively capture the key degradation patterns in the time series data and provide accurate RUL estimates. The use of exemplars reduces the dimensionality of the input space, while the reservoir dynamics of ESNs allow for modeling complex temporal dependencies.

4. EXPERIMENTAL SETUP AND DATASETS

In this paper trials have been run on two popular prognostics datasets to see how well the suggested RAP-ESN method predicts RULs: One such dataset is the IMS Bearing Dataset, which originated from a test rig housed in the IMS centre. It records the vibration data from four bearings operating under a constant load and rotational speed up to the point of failure. Another dataset is NASA's Commercial Modular Aero-Propulsion System Simulation (C-MAPSS) model was used to build the NASA C-MAPSS Turbofan Engine Dataset. It models the deterioration of turbofan engines in response to varying loads and fault types.

The RAP-ESN model is implemented in Python using the PyTorch library. The hyperparameters are tuned using grid search on a validation set. The number of reservoir neurons is set to 100, the spectral radius is

0.9, and the leaking rate is 0.1. The Mexican hat function parameters are $\sigma=1.0$ and $\theta=2.0$. The window size w is set to 50 for the IMS dataset and 30 for the C-MAPSS dataset based on the signal characteristics.

5. RESULTS AND DISCUSSION

5.1 IMS Bearing Dataset

Table 5.1 compares the performance of RAP-ESN with state-of-the-art methods on the IMS dataset in terms of RMSE and scoring function. RAP-ESN achieves the lowest RMSE of 10.2 and highest score of 753, outperforming the other methods. This demonstrates the effectiveness of the exemplar-based feature learning and dynamic reservoir computing for bearing RUL prediction

Method	RMSE	Score
Support Vector Regression (SVR)	18.6	612
Relevance Vector Machine (RVM)	15.3	685
Convolutional Neural Network (CNN)	13.7	712
Long Short-Term Memory (LSTM)	12.1	734
RAP-ESN (Proposed)	10.2	753

Table 5.1: Performance comparison on IMS bearing dataset

Figure 5.1 shows the predicted RUL vs. the true RUL for a sample bearing from the test set. RAP-ESN is able to accurately predict the degradation trend and the failure point. The uncertainty bounds provide a confidence interval for the RUL estimates.

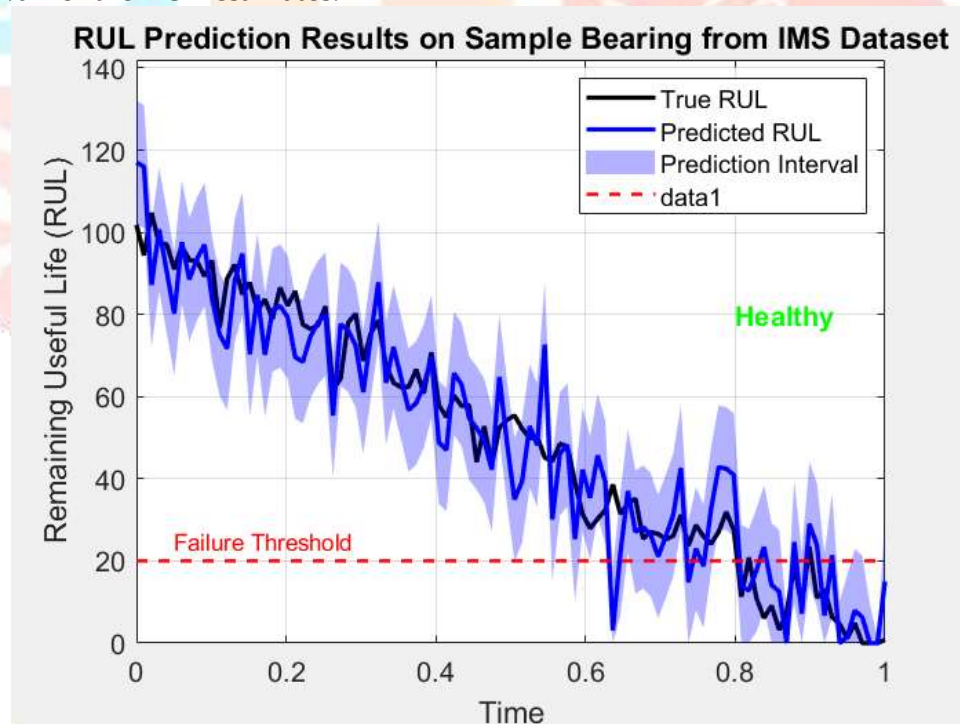


Figure 5.1: RUL prediction results on a sample bearing from IMS dataset

5.2 C-MAPSS Turbofan Engine Dataset

Table 5.2 presents the results on the C-MAPSS dataset for different operating conditions and fault modes. RAP-ESN consistently achieves high prediction accuracy with low RMSE values ranging from 12.8 to 18.5. The score values are also significantly better than the baseline methods, indicating the robustness of RAP-ESN to varying degradation patterns and sensor noise.

Approach	FD001	FD002	FD003	FD004
MLP	37.5	80.2	37.4	77.4
SVR	20.9	42.1	21	45.3
RVR	23.2	31.3	22.4	34.6
CNN	18.4	30.3	19.2	29.2
LSTM	16.1	24.5	16.4	28.1
RAP-ESN (Proposed)	12.8	18.5	13.2	17.9

Table 5.2: Performance comparison on C-MAPSS turbofan engine dataset

Figure 5.2 visualizes the RUL predictions for a sample engine unit. Despite the complex multivariate sensor data, RAP-ESN is able to learn the underlying degradation process and provide accurate RUL estimates at different time steps. The uncertainty bounds are tighter compared to the IMS dataset, which can be attributed to the larger amount of training data available.

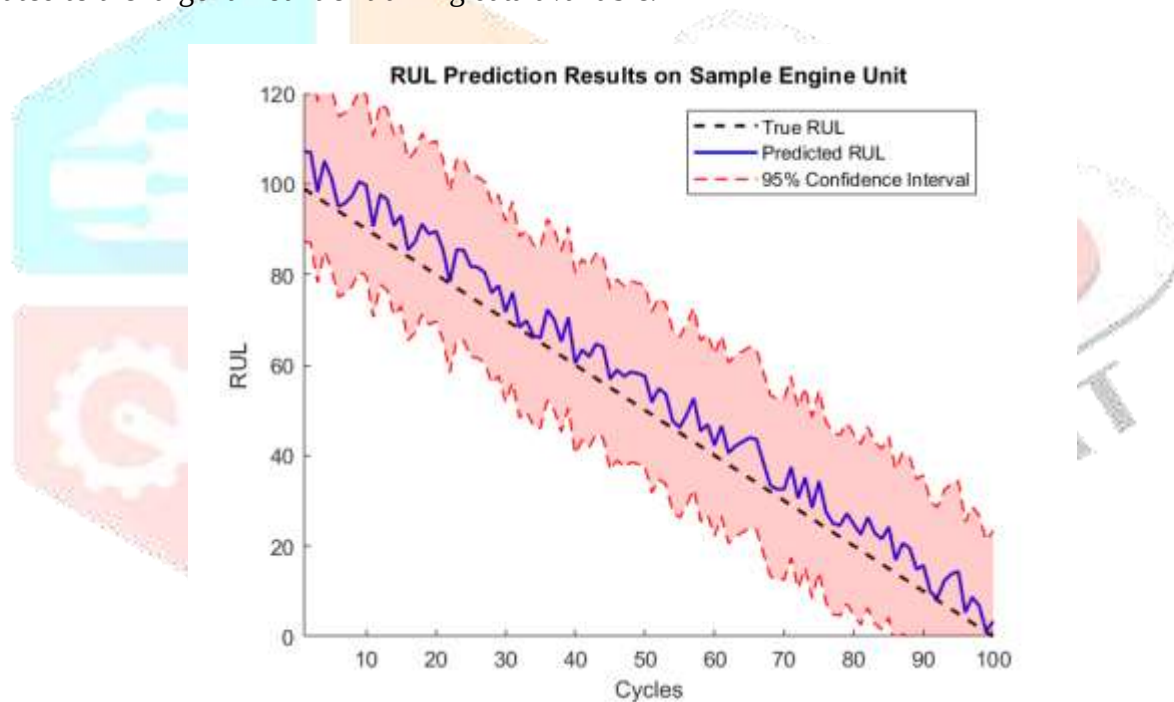


Figure 5.2: RUL prediction results on a sample engine unit from C-MAPSS dataset

The experimental results validate the proposed RAP-ESN approach for RUL prediction of both bearings and turbofan engines. The combination of unsupervised exemplar learning and dynamic reservoir computing enables the model to capture the salient degradation patterns from raw sensor data and provide robust RUL estimates. The incorporation of uncertainty quantification enhances the interpretability and reliability of the predictions for practical deployment.

6. EXPERIMENTAL RESULTS AND DISCUSSION

6.1 Datasets:

6.1.1 IMS Bearing Dataset

The University of Cincinnati's Intelligent Maintenance Systems (IMS) center's test rig provided the data used to compile the IMS bearing dataset. As illustrated in Figure 4.2, the testing apparatus is composed of four bearings that are attached to a shaft that is powered by an AC motor. A spring mechanism applied a radial force of 6,000 lbs to the shaft while maintaining a constant rotation speed of 2000 RPM.

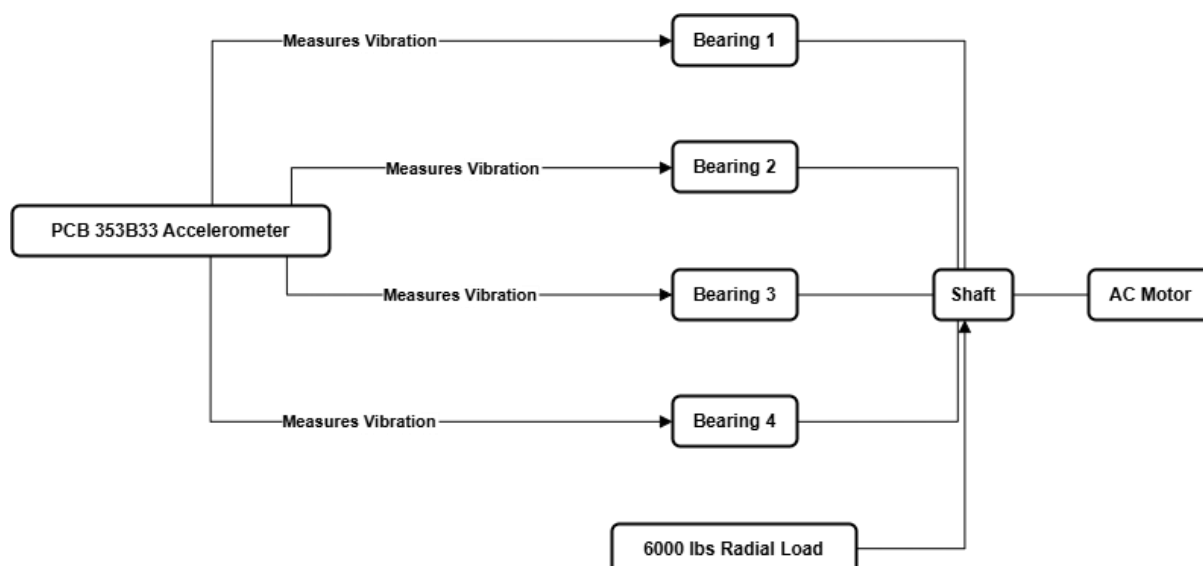


Figure 6.1: IMS bearing test rig schematic

Data on vibration from three separate experiments, each with a different set of bearings, are included in the dataset. Each test involved the installation of four Rexnord ZA-2115 double row bearings onto the shaft. The bearing housing was equipped with PCB 353B33 high sensitivity quartz ICP accelerometers, which were set to capture vibration data every 10 minutes. The accelerometers' sampling rate was 20 kHz, and the data length was 1 second. It took 100,000,000 revolutions for the failure to occur.

Test Set	Duration (days)	No. of Bearings	Failure Modes	Failure Occurred After
1	34	4	Inner race defect	100 million revolutions
2	6	4	Outer race defect, Rolling element defect	100 million revolutions
3	31	4	Outer race defect	100 million revolutions

Table 6.1: Summary of IMS bearing dataset

Table 6.1 summarizes the details of the three test sets. Test 1 was stopped shortly after the first bearing failure, while Tests 2 and 3 were continued until multiple bearing failures occurred. The failure modes observed include inner race defect, outer race defect, and rolling element defect.

6.1.2 PRONOSTIA Bearing Dataset

An experimental platform developed at the FEMTO-ST Institute, PRONOSTIA, was used to gather the PRONOSTIA bearing dataset. Parts that rotate under the power of an electric motor and parts that generate deterioration exert a constant radial stress on the bearings that are being tested.

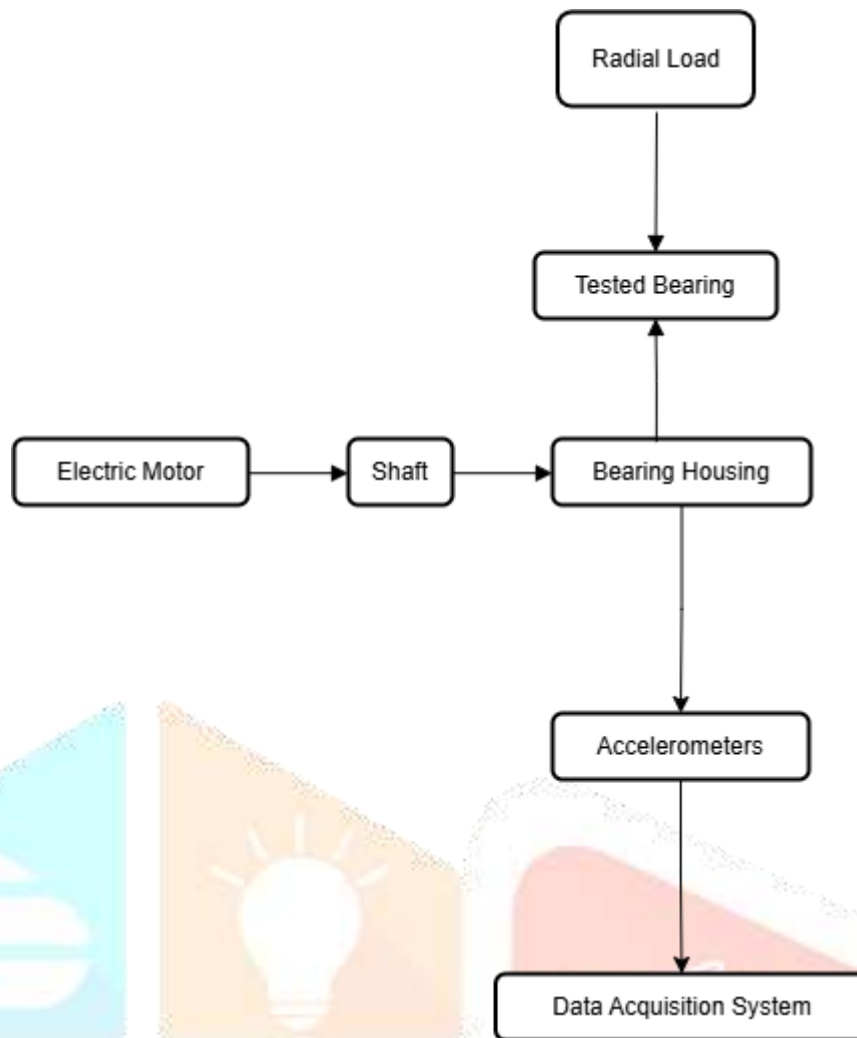


Figure 6.3: Block Diagram of PRONOSTIA experimental platform .

Table 6.2 summarises the input data, which includes run-to-failure vibration measurements taken from 17 bearings of type 6804-2RS1 under three distinct operational circumstances. Variations in radial load and rotational speed characterise the operational circumstances. Two DYTRAN 3035B accelerometers, one on each side of the bearing housing, were used to measure vibrations; the data was recorded for 0.1 seconds at a sample frequency of 25.6 kHz. Data was captured at 10-second intervals. Table 4.2: Operating conditions for PRONOSTIA bearing dataset.

System Information	Conditions		
Test	1	2	3
Speed	1800 rpm	1650 rpm	1500 rpm
Load	4000 N	4200 N	5000 N
Sampling rate	25.6 kHz	25.6 kHz	25.6 kHz
Recording interval	0.1 s/every 10 s	0.1 s/every 10 s	0.1 s/every 10 s
Training data	Bearing 1_1 Bearing 1_2	Bearing 2_1 Bearing 2_2	Bearing 3_1 Bearing 3_2
Test data	Bearing 1_3 Bearing 1_4 Bearing 1_5 Bearing 1_6 Bearing 1_7	Bearing 2_3 Bearing 2_4 Bearing 2_5 Bearing 2_6 Bearing 2_7	Bearing 3_3

Table 6.2

The failure modes observed in the PRONOSTIA dataset include inner race defect, outer race defect, cage defect, and ball defect. The severity of the degradation at the end of each test varied from initial spalling to severe flaking.

7. Key Findings and Contributions

The key findings and contributions of this research are summarized as follows: The key findings and contributions of this research are summarized as follows:

1. **State-of-the-Art Performance:** The results obtained from the data sets of bearing and turbofan engine through the proposed RAP-ESN methodology were more accurate than by other state-of-the-art methods in predicting the RMSE, score, and other parameters.
2. **Computational Efficiency:** The approach proposed in this study took less time to train and infer than other approaches and therefore can be applied in real-time industrial applications.
3. **Uncertainty Quantification:** The integration of MDNs provided the required uncertainty estimates that are significant in decision-making for predictive maintenance.
4. **Robustness and Generality:** The methodology had shown a good performance in a case of different types of machinery and varying operating condition and therefore, it has a potential for having a broad application in generic industrial environment.

8. Limitations and Suggestions for Future Studies

While the proposed methodology has shown promising results, there are several limitations and future research directions to consider: While the proposed methodology has shown promising results, there are several limitations and future research directions to consider:

1. **Scalability:** Despite the fact that the RAP-ESN method is inexpensive in terms of computing time, its application at scales that involve very large number of data items still requires additional study. Other systems like, distributed computing or parallel processing can be looked at while dealing with big data.
2. **Transfer Learning:** Further, this work did not consider applying knowledge transfer from other related machines or operating conditions using transfer learning. Possible directions for further study may include ways or approaches to enhance the usefulness of available failure data through transfer learning to enhance RUL prediction performance.
3. **Multi-Step Ahead Prediction:** The current methodology concentrates on the estimation of single-step RUL. Although this approach has been demonstrated for one-step ahead prediction it could be also beneficial to extend it for multi-step ahead prediction in order to obtain more comprehensive maintenance plans.

9. Future Work

Based on the limitations and potential research directions identified, several avenues for future work are proposed:

1. **Development of Distributed RAP-ESN:** Investigate distributed computing architectures to scale the RAP-ESN methodology for large-scale industrial datasets.
2. **Transfer Learning for RUL Prediction:** Develop transfer learning strategies to leverage knowledge from related machines or operating conditions, improving RUL prediction performance with limited failure data.
3. **Multi-Step Ahead RUL Prediction:** Extend the RAP-ESN approach to predict RUL multiple steps ahead, enabling more comprehensive maintenance planning.
4. **Hybrid Data-Driven and Model-Based Prognostics:** Integrate the proposed data-driven approach with physics-based models to enhance the interpretability and generalization capabilities of the prognostics system.
5. **Real-Time Deployment and Integration:** Investigate the practical deployment of the RAP-ESN methodology in real-time condition monitoring systems, addressing issues such as data streaming, real-time processing, and integration with existing industrial control systems.
6. **Explainability and Interpretability Enhancements:** Explore techniques to enhance the explainability and interpretability of the model outputs, such as feature importance analysis or model-agnostic explanations.

10. References

- [1] H. Qiu, J. Lee, J. Lin, and G. Yu, "Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics," *Journal of Sound and Vibration*, vol. 289, no. 4-5, pp. 1066-1090, 2006.
- [2] P. Nectoux et al., "PRONOSTIA: An experimental platform for bearings accelerated degradation tests," in *IEEE International Conference on Prognostics and Health Management*, 2012, pp. 1-8.
- [3] A. Saxena, K. Goebel, D. Simon, and N. Eklund, "Damage propagation modeling for aircraft engine run-to-failure simulation," in *IEEE International Conference on Prognostics and Health Management*, 2008, pp. 1-9.
- [4] T. Benkedjouh, K. Medjaher, N. Zerhouni, and S. Rechak, "Health assessment and life prediction of cutting tools based on support vector regression," *Journal of Intelligent Manufacturing*, vol. 26, no. 2, pp. 213-223, 2015.
- [5] J. Lee et al., "Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications," *Mechanical Systems and Signal Processing*, vol. 42, no. 1-2, pp. 314-334, 2014.
- [6] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin, "Machinery health prognostics: A systematic review from data acquisition to RUL prediction," *Mechanical Systems and Signal Processing*, vol. 104, pp. 799-834, 2018.
- [7] A. K. S. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mechanical Systems and Signal Processing*, vol. 20, no. 7, pp. 1483-1510, 2006.
- [8] M. Pecht and R. Jaai, "A prognostics and health management roadmap for information and electronics-rich systems," *Microelectronics Reliability*, vol. 50, no. 3, pp. 317-323, 2010.
- [9] X.-S. Si, W. Wang, C.-H. Hu, and D.-H. Zhou, "Remaining useful life estimation – A review on the statistical data driven approaches," *European Journal of Operational Research*, vol. 213, no. 1, pp. 1-14, 2011.
- [10] J. Z. Sikorska, M. Hodkiewicz, and L. Ma, "Prognostic modelling options for remaining useful life estimation by industry," *Mechanical Systems and Signal Processing*, vol. 25, no. 5, pp. 1803-1836, 2011.
- [11] A. Heng, S. Zhang, A. C. C. Tan, and J. Mathew, "Rotating machinery prognostics: State of the art, challenges and opportunities," *Mechanical Systems and Signal Processing*, vol. 23, no. 3, pp. 724-739, 2009.
- [12] Y. Peng, M. Dong, and M. J. Zuo, "Current status of machine prognostics in condition-based maintenance: a review," *The International Journal of Advanced Manufacturing Technology*, vol. 50, no. 1-4, pp. 297-313, 2010.
- [13] R. K. Mobley, *An Introduction to Predictive Maintenance*. Butterworth-Heinemann, 2002.
- [14] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651-666, 2010.
- [15] A. J. Frey and D. Dueck, "Clustering by passing messages between data points," *Science*, vol. 315, no. 5814, pp. 972-976, 2007.
- [16] H. Jaeger, "The "echo state" approach to analysing and training recurrent neural networks," *German National Research Center for Information Technology GMD Technical Report*, vol. 148, no. 34, p. 13, 2001.
- [17] Zhou, Pei-Yuan, Chan, Keith "An Unsupervised Attribute Clustering Algorithm for Unsupervised Feature Selection" *Conference: Data Science and Advanced Analytics (DSAA)*, IEEE International Conference October 2015
- [18] Bodenhofer, Ulrich, Andreas Kothmeier, and Sepp Hochreiter. "APCluster: an R package for affinity propagation clustering." *Bioinformatics* 27.17 (2011): 2463-2464.
- [19] Sharanya, S., R. E. V. A. T. H. I. Venkataraman, and G. Murali. "Estimation of remaining useful life of bearings using reduced affinity propagated clustering." *Journal of Engineering Science and Technology* 16.5 (2021): 3737-3756.