

A Systematic Review of Remaining Useful Life Prediction in Roller Bearings Using Artificial Intelligence Techniques

Rajkumar Palaniappan

Department of Mechatronics Engineering, College of Engineering, University of Technology Bahrain, Kingdom of Bahrain

E-mail: prkmect@gmail.com

Keywords: RUL ,artificial intelligence, prediction, predictive maintenance, literature, review

Received: October 6, 2025

Artificial Intelligence (AI) has demonstrated to be an effective method for predicting irregularities across various Industrial processes. In recent years, the development of predictive maintenance systems using AI techniques has attracted many researchers around the world. Maintenance planning has been effective with Remaining useful life prediction using AI in bearings. This review brings consideration to the role of AI in predicting the remaining useful life (RUL) of bearing components in various industrial processes. A systematic search was carried out in electronic databases such as Springer, IEEE, Elsevier, and the ACM Digital Library, with an emphasis on AI-based approaches for bearing RUL prediction. A brief summary of previous works is presented to show the development of technological advancement in this field of RUL prediction of roller bearing. Specifically, this review examines the types of bearing components studied, the sample sizes used for training AI models, the signal processing method and classification algorithm applied, and the outcomes achieved. The outcome of this review shows that hybrid approaches and deep learning models achieve better performance in predicting RUL in roller bearings. Finally, the review finds existing research gaps and provides recommendations for future improvements, aiming to guide future researchers toward more accurate and reliable RUL prediction models for bearings.

Povzetek: Pregled sistematično povzema raziskave o uporabi AI za napoved preostale življenjske dobe (RUL) valjčnih ležajev, pri čemer ugotavlja, da hibridni pristopi in globoko učenje praviloma dosegajo najboljše rezultate, hkrati pa izpostavi raziskovalne vrzeli in priporočila za razvoj natančnejših ter zanesljivejših modelov.

1 Introduction

An important technique in predictive maintenance is Remaining Useful Life (RUL), sometimes stated as Remaining Useful Life. It involves classifying or predicting how long a particular machine or component can be used before deteriorating completely or reaching a threshold [1]. An important phase of maintenance planning is the RUL estimation which improves equipment performance and reliability overall, optimizes maintenance strategies, minimizes downtime and lowers operational costs [2], [3]. Predicting or classifying a component or machine's RUL is mainly done to proactively schedule maintenance tasks according to the component's/ machines anticipated RUL. Maintenance responsibilities can be planned and done ahead of time by accurately estimating the RUL, ensuring that any necessary actions are taken before failures occur [4], [5]. This approach reduces the possibility of disastrous failures, maximizes resource use, and avoids costly machine downtime. Knowing the degradation state of a component or machine and being able to monitor its health are necessary for predicting its RUL [6]. The term "degradation state" describes how a component or a machine performance or health status changes over time,

indicating how long it will continue to function properly or in other words how long a component or a machine function before it fails [7], [8]. Analyzing historical data, detecting anomalies in sensor data, or researching how similar components behave under constant operating conditions can all be used to describe these patterns. Monitoring a component or machine's health involves gathering relevant data using various sensors, including vibration, temperature, acoustic, and oil quality sensors [9], [10]. Important parameters that show the component's performance and overall health are continuously measured using these sensors. It is possible to track the degradation process and identify deviations from standard operating conditions by monitoring these parameters. Several data driven algorithms are used to accurately estimate the RUL [11], [12]. Regression models, neural networks, decision trees, and support vector machines are a few examples of machine learning algorithms that are often used to examine historical data, identify anomalies, and forecast the RUL [13], [14]. These algorithms establish connections between the degradation patterns and the anticipated RUL by using features or variables that are extracted from the sensor data. There are no specific methods or algorithms that can be used for RUL estimation. Data preprocessing, feature selection, model training, and validation are frequently employed

methodology [15]. The accuracy of RUL predictions can be impacted by several factors such as the availability and quality of data, the selection of relevant features, the selection of suitable degradation models, and the assessment of uncertainties and variability. Real time RUL estimation is now possible due to technological advancements like edge computing and the Internet of Things (IoT) [2], [16]. It is possible to gather real time data, provide RUL predictions, and continuously monitor the health condition of components or machines by using edge computing techniques and integrating sensor devices with IOT capabilities [17]. Condition-based maintenance decisions, quick maintenance interventions, and improved component or machine management efficiency is all made possible by real time RUL estimation [3], [7]. To sum up, estimating Remaining Useful Life (RUL) is an important technique in predictive maintenance. It involves predicting a machine's or component's RUL based on health condition monitoring and degradation patterns. Proactive maintenance planning is made possible by accurate RUL estimation, which also improves equipment performance and reliability overall, minimizes downtime, and lowers reactive maintenance expenses. Industries can optimize maintenance plans and ensure the efficient use of resources by employing data driven techniques and algorithms.

2 Methodology

This systematic review follows the PRISMA guidelines to ensure a clear and thorough approach to our research. At first, all related articles were identified through searches on various electronic databases, including Springer, IEEE, Elsevier, and ACM Digital Library, covering literature published from 2008 to 2025. The initial search applied the keywords "Remaining Useful Life Prediction," yielding many articles. A successive search was conducted using the keywords "Artificial Intelligence" and "Bearing," which resulted in 153 articles. A selection criterion was established, and articles were chosen based on the following criteria: (i) relevance to remaining useful life prediction, (ii) inclusion of artificial intelligence techniques in predictive maintenance, (iii) publication in English, (iv) exclusion of duplicate articles identified through title and author matching, (v) focus on bearing remaining useful life prediction, and (vi) exclusion of

articles highlighting predictive maintenance without considering artificial intelligence. This process resulted in a total of 153 articles from the initial search. Following a review of titles and abstracts, 92 articles were excluded, and the remaining articles were examined in detail, leading to the exclusion of 29 additional articles due to insufficient information. Ultimately, 32 articles that met all the selection criteria were included in the final selection. This number was deemed sufficient to provide a comprehensive overview of the current state of research in this field, based on the saturation of themes. Data extraction focused on key performance metrics such as RMSE and accuracy, which were normalized where applicable to facilitate fair comparisons. Figure 1 illustrates the flowchart of the selection criteria.

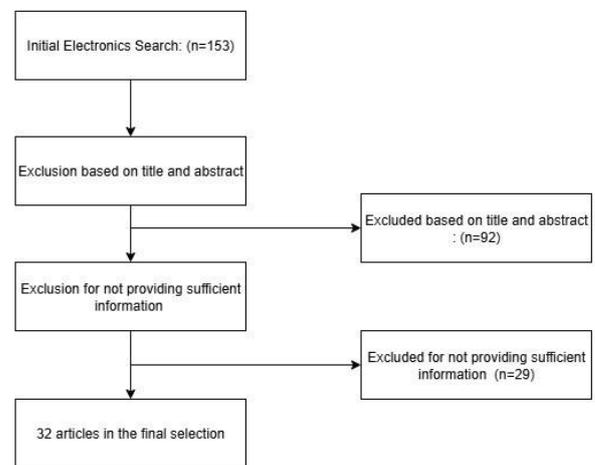


Figure1: Selection criteria

3 Overview of the literature search

A summary of the 31 articles that met the selection criteria is tabulated in Table 1. This overview highlights the dataset used, preprocessing, feature engineering techniques and machine learning techniques utilized by previous researchers in predicting remaining useful life prediction of bearing components.

Table 1: Recent works related to RUL prediction in bearing using artificial intelligence

| Reference | Bearing type | Dataset | Preprocessing/ Feature Extraction | Methodology | Outcome |
|-----------|--------------|------------------|--|---|---|
| [18] | Ball Bearing | Vibration signal | Root Mean Square (RMS), Kurtosis, Crest Factor | Self-organizing map in combination with backpropagation model | Nearly 85% of the predictions from the proposed model fall within a 20% margin of error of the actual bearing failure time. |

| | | | | | |
|------|-------------------------------------|---|---|--|--|
| [19] | Taper roller bearings | electrostatic wear-site and vibration data | RMS data | Gaussian Mixture Model (GMM) | Satisfactory results were obtained |
| [20] | FAG 7206 B single ball bearings | Vibration data | discrete wavelet transform (DWT) | multilayer perceptron (MLP) network | Classification accuracy of around 85% was reported |
| [21] | Bearing | Vibration signal | Principle component analysis (PCA) | Support Vector Machine (SVM) | errors are less than 4%, and the average error is 3.2%. |
| [22] | Bearing | Vibration data (128 samples for various stages of degradation) | mean, rms, shape factor, skewness, kurtosis, crest factor, entropy estimation, entropy estimation error, histogram lower and upper. | Support Vector Machine (SVM) | Average prediction accuracy was 94.4% |
| [23] | Bearing | Historical data comprising of age values and actual condition monitoring measurement values | Generalized Weibull-FR function | Artificial Neural Network (ANN) | Average prediction error of 10.6 % was obtained |
| [24] | NSK 6804DD bearing | Speed and Load variation | temporal, frequency, or time-frequency features | Mixture of Gaussians Hidden Markov Models | Accuracy of 0.7874 reported |
| [25] | Bearing | Vibration data | Time-domain feature | support vector machine | Root Mean Square Error (RMSE) = 2.2353 |
| [26] | Bearing NSK 6804RS | Vibration data from 3 bearings | Hilbert-Huang transform (HHT) | support vector machine | Error ranged from 0.6 to 1.25% |
| [27] | Rexnord ZA-2115 double row bearings | Vibration measurement using accelerometers. Load and speed were also recorded | RMS and Kurtosis Features fitted using Weibull distribution | simplified Fuzzy adaptive resonance theory map neural network (SFAM) | 1 normal class and 6 states of degradation were predicted. With almost 99 % classification accuracy for normal class and with reduction of classification accuracy for the other states. |
| [28] | Gearbox Bearing | Temperature data 10 minutes data | Raw signal | ANN | RMSE below 1 reported |
| [29] | ball bearing (SKF 51210) | piezoelectric sensor (physical acoustic corporation type "PICO") sensor data (totally 4 sensors used) | Acoustic Emission (AE) feature extraction | ANN | Satisfactory results were obtained |

| | | | | | |
|------|-------------------------------|--|--|---|---|
| [30] | Bearing | Vibration signal | Time domain features - root mean square, crest factor, and kurtosis Frequency domain frequency spectrum | Deep neural network | RMSE achieved close to 0. |
| [31] | Bearing | Vibration signal | Restricted Boltzman Machine to automatically extract features | deep learning-based DBN-FNN | RMSE = 2.64 |
| [32] | Bearing | IEEE PHM2012 Predictor Challenge experiment data (Vibration signal) | spectrum-principal-energy-vector | Deep Convolution Neural Network | RMSE of 0.119 was obtained |
| [33] | Roller Bearing | NASA FEMTO BEARING DATASET | - | Neural Network regression (NNR) | Predicted score for NNR is 5,537 secs |
| [9] | Bearing | force lubrication, constant speed and load - IMS dataset and PRONOSTIA dataset | Relative Root Mean Square (RRMS) and Inertial Relative Root Mean Square (IRRMS) | SVM | 98.88% Classification accuracy was obtained |
| [34] | rolling element bearings | PRONOSTIA | Degradation indicator DEI extraction | Convolutional Neural Network (CNN) | NRMSE = 0.05 |
| [35] | Bearing | IEEE PHM 2012 PRONOSTIA dataset | Statistical Feature - RMS values | generative adversarial network (GAN) architecture combining the LSTM network and autoencoder (AE) | NRMSE = 0.0456 |
| [36] | Bearing | NASA IMS dataset | 3 sigma criterion applied to remove noise and FFT applied followed by root mean square | deep convolutional neural network | RMSE = 0.01818 and $R^2 = 0.95846$ |
| [37] | Bearing | Vibration signal (University of California Irvine dataset) | time domain, frequency domain, and entropy features | Weighted K-Nearest Neighbor (K-NN) | Almost 100 % accurate |
| [38] | Roller Bearing | PRONOSTIA test set | Fuzzy logic | CNN | RMSEs was 20.24% |
| [13] | Roller bearing | Roller bearing dataset | Principal component analysis | KNN | MAPE, MSE, RMSE 95% confidence in predicting the RUL |
| [39] | rolling element bearing (REB) | Vibration signals, 10,000 data | continuous Wavelet transform (CWT) and extraction of statistical features from CWT coefficients. | KNN | 83.3% Classification accuracy |
| [40] | Angular contact | Acoustic emission signal | Prognostic feature selection | Multilayer Perceptron (MLP) and | MLP- Squared error = 7.86 RBF - |

| | | | | | |
|------|--|--|---|---|---|
| | ball bearing | | | Radial Basis function (RBF) neural networks | Mean Squared error = 2.85 |
| [41] | Roller Bearing | 6180 data points (temperature, speed, load, dimensions of the inner and outer rings, width, vibration amplitude, vibration frequency, lubricant type, and lubricant viscosity) | Principal component analysis (PCA) | SVM, KNN and Random Forest classifiers | The results reveal that Support Vector Machines topped the charts with an impressive 96.74% accuracy, while Random Forest came in second at 95.95%, and k-Nearest Neighbors followed with 91.77%. |
| [42] | bearings | Vibration data | stacked variational denoising autoencoder (SVDAE) | bidirectional long short-term memory (BiLSTM) neural network | Mean squared error = 2.2304%. |
| [43] | Bearing | Vibration signal | peak value, RMS, magnitude and spectral partition sum | Deep Neural Network combined with knn | Maximum absolute error = 0.0920 |
| [44] | Ball Bearing | dataset of 2155 vibration signals | mean, standard deviation, kurtosis, and root mean square | Extreme Gradient Boost | 96.61% |
| [45] | Rolling bearing | Voltage signal | Short-Time Fourier Transform | Hybrid MLP and LSTM Neural Network | 99% accurate |
| [46] | Ball Bearing (Class 6204 2RS bearings) | Vibration signal (27,725 records) of normal and faulty conditions | Statistical and frequency-domain features followed by PCA | Hybrid SVM and ANN | 95 % accurate |
| [47] | Ball Bearing | Vibration data (Two datasets were used) | global topology aggregation | Dual-Channel Dynamic Spline Graph Convolutional Network (DDSGNet) | Better prediction accuracy compared to conventional methods were reported |

The studies on RUL prediction in bearing can be classified into three phases namely, studies before 2012, studies between 2012 and 2017 and studies after 2017. Figure 2 shows the model usage across phases in predicting Rul in roller bearings. The initial studies in prediction of RUL for bearings relied profoundly on statistical features and classical machine learning models. [18] extracted features such as RMS, kurtosis, and crest factor from vibration signals and used a self-organizing map with backpropagation algorithm in classifying RUL of bearing

and achieved nearly 85% of predictions within a 20% error margin. Similarly, [19] used Gaussian Mixture Models (GMM) in classifying Rul of bearing with RMS data collected from taper roller bearings and reported satisfactory results. [20] employed discrete wavelet transform (DWT) based features and utilized multilayer perceptron networking in classifying RUL of bearing and achieved approximately 85% classification accuracy. [21] combined PCA-based dimensionality reduction with SVM, achieving an average error of just 3.2%.

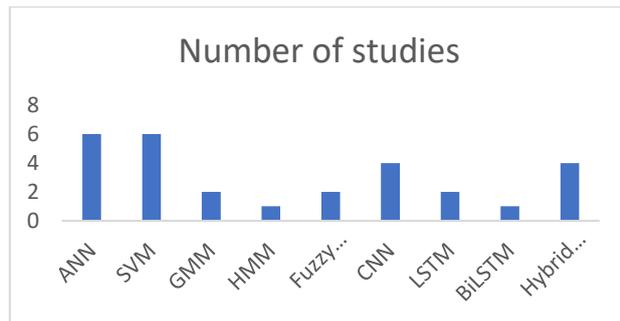


Figure 2: Model usage in bearing RUL studies

Between the year 2012 to 2017, research began to use more sophisticated statistical and time–frequency features with improved machine learning algorithms. [22] extracted a diverse set of statistical features such as mean, RMS, skewness, kurtosis, entropy, crest factor and applied SVM classifier in classifying RUL in bearings and achieved 94.4% accuracy. [23] modeled the degradation pattern in bearing with an ANN model combined with Weibull-FR functions and reported an average prediction error of 10.6%. [24] combined Gaussians within a Hidden Markov Model and achieved 78.7% accuracy. [25] applied time-domain features with SVM classifier in classifying RUL in bearing which yielded an RMSE of 2.2353. Other notable contributions included Hilbert–Huang transform based feature extraction with SVM Classifier [26], error 0.6–1.25%) and fuzzy ART neural networks [27], which reached approximately 99% classification accuracy for normal states but lower performance for degradation states. These studies improved accuracy compared to earlier studies, up to 90 to 95% accuracy but still depended on careful feature engineering and were susceptible to noise and unpredictable operating conditions.

Since 2017, the field has been dominated by advancement in deep learning techniques which has reduced the dependence on custom extracted features. [30] combined feature extraction techniques from time- and frequency-domain along with deep learning approach and achieved RMSE values close to zero. [31] utilized deep belief networks with restricted Boltzmann machines for feature extraction and reported RMSE value of 2.64. [32] used deep CNNs on PHM2012 dataset and obtained RMSE value of 0.119. [34] applied CNNs to PRONOSTIA dataset and reported NRMSE to be 0.05, while [36] achieved RMSE value of 0.01818 and $R^2 = 0.95846$ using FFT preprocessing followed by a deep CNN technique.

Generative and hybrid approaches further advanced the field of RUL prediction in bearing further. [35] developed a GAN–LSTM–autoencoder model and tested it using PHM2012 dataset. They obtained NRMSE value of 0.0456. [37] applied weighted KNN with time domain, frequency domain, and entropy features, obtaining nearly 100% accuracy. [41] incorporated multimodal inputs such as temperature, speed, load, dimensions, and lubricant properties and PCA for dimension reduction. They reported 96.74% with SVM classifier, 95.95% with

Random Forest classifier, and 91.77% with KNN classifier. Many recent approaches have used denoising autoencoders with BiLSTM [42] and deep neural networks combined with KNN [43], reporting less mean squared error compared to other studies. Ensemble learning methods also remain effective in RUL prediction: [44] used XGBoost to achieve 96.61% accuracy, while [45] reported 99% accuracy with a hybrid MLP–LSTM trained. This phase has consistently delivered advancement, with most deep learning or hybrid methods exceeding 95% accuracy or reporting nearly zero RMSE on standard benchmarked datasets. In all the phases mentioned above, public open-source benchmark datasets have played a central role. PHM2012 has been widely used in deep learning works, such as [32] and [35]. PRONOSTIA has also been a main choice, as seen in [33], [34], and again [35]. The NASA IMS dataset was used in the work of [36], while the FEMTO-ST dataset was used in the work of [33]. These datasets provide controlled run-to-failure experimental data, enabling impartial method comparison. However, the constant load and speed limit generalization to real industrial conditions. To address this, recent works have adapted multimodal or alternative sensing data for RUL predictions in roller bearing. For example, [41] combined vibration with temperature, load, lubricant properties, and geometric parameters, while Motahari-[40] used acoustic emission signals both showing improved robustness beyond vibration inputs alone. Preprocessing phase remains vital even in the deep learning era. Conventional signal processing techniques are still widely applied: FFT has been used by [36], discrete wavelet transform (DWT) has been used by [20], and continuous wavelet transform (CWT) has been used by Kumar & [39]. Principal Component Analysis (PCA) has been employed for dimensionality reduction in the works of [21], [13], and [41]. More recently, denoising autoencoders have marked a shift toward machine learning based preprocessing approaches, with [42] applying a stacked variational denoising autoencoder with BiLSTM to adaptively reduce noise and improve the feature quality.

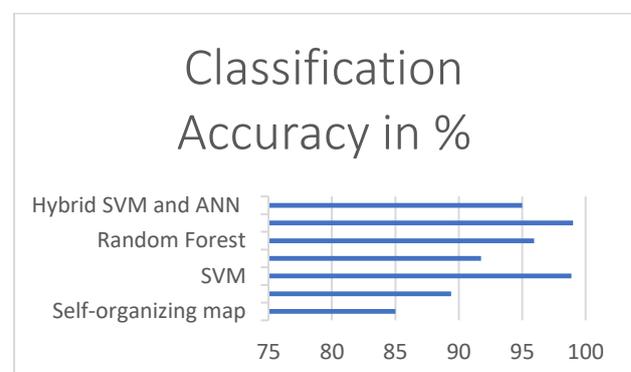


Figure 3: Bar chart comparing the performance outcome in terms of classification accuracy

Figure 3 shows the classification accuracy reported by using various classification models in predicting roller bearing RUL. The comparison of outcomes from recent works on Remaining Useful Life (RUL) prediction for bearings using artificial intelligence reveals noteworthy distinctions in performance across various methodologies. The Support Vector Machine (SVM) model emerged as a frontrunner, achieving a remarkable classification accuracy of 96.74% (Reference [26]), indicating its robust capacity to manage complex datasets effectively. In contrast, the k-Nearest Neighbor (KNN) classifier reported a lower accuracy of 91.77%, suggesting that while KNN is a valuable tool, it may not match the predictive power of SVM in more intricate scenarios. Deep learning approaches also showed compelling results, with one model reporting a root mean square error (RMSE) as low as 0.01818 (Reference [20]). This highlights the precision that deep learning techniques can bring to RUL predictions. However, other models employing hybrid methodologies illustrated variability, with RMSE values ranging from 0.05 to 2.64, reflecting both the effectiveness and challenges involved in optimizing these advanced systems. Feature extraction methods, such as Principal Component Analysis (PCA), demonstrated strong performance, achieving RMSE values below 1 (Reference [11]). This indicates that proper data reduction techniques can significantly enhance the accuracy of predictions. Conversely, techniques like Gaussian Mixture Models (GMM) yielded satisfactory outcomes but lacked comprehensive metrics for direct comparison, pointing to a need for clearer measurement standards. Interestingly, traditional statistical methods also reported competitive outcomes, with accuracies between 83.3% and nearly 100% across specific applications, such as those using continuous wavelet transforms combined with KNN (Reference [24]). This suggests that while advanced methods dominate the discourse, simpler statistical techniques can still deliver reliable results, especially in less complex environments. In summary, the comparison of outcomes across these studies underscores the diversity of methodologies available for RUL prediction. While SVM and deep learning techniques consistently provide high accuracy, traditional methods and hybrid approaches also remain relevant, making a strong case for incorporating a variety of strategies to optimize predictive maintenance in engineering applications.

Models developed for predicting the remaining useful life (RUL) of mechanical systems share many conceptual similarities with adaptive and robust control methods. Both aim to maintain reliable operation when the system is affected by uncertainty, nonlinearity, and changing conditions. In adaptive control, the controller continuously adjusts its parameters based on feedback to preserve system stability, even when the underlying dynamics are not perfectly known. In a similar way, data-driven RUL models learn degradation patterns from sensor measurements and update their understanding of

system health as new information becomes available. The two approaches are built on the same principles of learning, adaptation, and robustness, although they address different time scales: adaptive control manages short-term stability, while RUL prediction focuses on long-term performance and reliability. Research on adaptive fuzzy control and robust neural adaptive control has shown how intelligent models can compensate for nonlinear behavior and unpredictable disturbances. The same ideas are reflected in deep learning-based RUL models, which approximate complex degradation processes while remaining resilient to noise and missing data. Likewise, adaptive backstepping control methods used for nonlinear and flexible robotic systems share structural similarities with layered deep learning architectures such as CNN-LSTM and BiLSTM models, where each stage refines information and reduces uncertainty to improve prediction accuracy and stability.

Bringing these two domains together offers the possibility of turning predictive maintenance into an active and self-correcting control process. Concepts from nonlinear optimal control and model-based predictive control closely mirror the goals of RUL estimation, where optimization seeks to minimize energy use, error, or component wear based on the predicted health of the system. In an integrated framework, the RUL predictor acts as a health observer that anticipates future degradation, while the adaptive controller serves as the decision-maker, adjusting parameters such as torque, load, or speed to prevent further damage and extend service life. This interaction creates a closed feedback loop in which prediction and adaptation work together to sustain performance under uncertain conditions. Embedding control-theoretic stability ideas, such as Lyapunov convergence, within RUL modeling can also provide a stronger theoretical foundation, allowing models to be assessed not only for accuracy but also for stability and reliability. Combining predictive modeling with adaptive control therefore points toward the next generation of intelligent maintenance systems—machines that can monitor their own condition, adapt their operation in real time, and continue functioning safely even in the presence of faults or changing environments.

4 Discussion

Remaining Useful Life (RUL) can be viewed in two ways: as a regression problem where we predict a continuous value, or as a classification problem when we break it down into specific categories. The decision to use one approach over the other often depends on the goals of the study and how researchers decide to define the problem. The journey of predicting the Remaining Useful Life (RUL) of bearings has transformed significantly, especially with the emergence of deep learning. In the early stages, researchers depended on statistical features and traditional machine learning methods, which often struggled with noise and changing conditions. From 2012

to 2017, techniques like Support Vector Machines (SVM) and Artificial Neural Networks (ANN) began to enhance accuracy by utilizing more complex features, though they still relied heavily on manual feature engineering. Today, deep learning allows us to create models that can learn directly from raw data, resulting in impressive outcomes. However, this advancement also makes it more challenging to understand how these models arrive at their predictions, raising concerns about interpretability. It's essential to establish a stronger theoretical framework by comparing these models with adaptive and robust control techniques like adaptive fuzzy control and robust neural adaptive control that effectively manage uncertainty in systems. Despite these advancements, several challenges remain. Many datasets lack the diversity needed, which limits how well our models can generalize to real-world situations. For instance, while some recent models perform well on benchmark datasets like PHM2012, they often struggle to transfer their effectiveness to real-world applications, where conditions can vary significantly. Future studies should incorporate cross-dataset evaluations to improve generalizability. Developing multimodal datasets that include various types of sensor data is vital for boosting robustness. Reproducibility is a concern due to missing hyperparameters and unclear preprocessing steps, so future studies should report details like model size, training epochs, and optimizer settings. We also need to standardize preprocessing techniques to improve reproducibility in research. Methods such as discrete wavelet transform (DWT) and Hilbert-Huang transform have been effective, but more comparative analyses are needed to determine the best preprocessing methods for different scenarios. Additionally, adopting probabilistic forecasting methods can help provide clearer insights into risks, facilitating better decision-making. Connecting RUL estimation to concepts like adaptive backstepping control and nonlinear optimal control can significantly enhance our predictive maintenance strategies. These approaches can aid in real-time decision-making and improve fault tolerance. For example, if a model predicts a potential failure, adaptive control can adjust operational parameters or schedule maintenance in advance, helping to reduce downtime and ensure system reliability. In practical terms, lightweight models that maintain accuracy are crucial for real-world applications. Techniques like model pruning and quantization can make these models more feasible in industrial environments. Moreover, it's important to address the risks of overfitting on benchmark datasets and to discuss the comparative weaknesses in interpretability or robustness to noise across different architectures. Additionally, we should thoroughly examine the challenges of deploying these models in real-world scenarios. In summary, while we've made great progress in RUL prediction, it's essential to tackle issues like generalization, preprocessing, and practical deployment. Additionally, integrating control strategies and ensuring that our models are interpretable

and robust will better position RUL prediction techniques to meet the needs of modern industry.

5 Future direction

Future progress in the field of RUL prediction for bearing depends on addressing these gaps completely. Foremost thing is the development dataset which includes multimodal sensor data with varying conditions. Collaborative efforts between academia and industry could lead to development of open-access datasets that capture variable speeds, loads, lubrication regimes, vibration, temperature and other environmental conditions. Extending the work of studies like [40] [41], which incorporated multiple sensor types, will be essential for building models that generalize beyond vibration-only signals. Such datasets should also be developed to overcome imbalance, ensure sufficient representation of various forms of failure that might occur. Secondly, the research of preprocessing and feature engineering should advance simultaneously. While deep learning reduces dependence on feature engineering, preprocessing is still important for filtering noise and enhancing the signal. Comparative analyses are needed to standardize conventional methods such as WPD and EMD against other alternatives such as denoising autoencoders and so on [37]. Preprocessing techniques implemented should be reported properly in publications, including denoising, normalization, windowing, and feature selection to improve reproducibility and allow fair cross study comparison. Thirdly, research on modeling should advance beyond testing with only one dataset. It's important to test models with different datasets, like training on one benchmark (e.g., PHM2012) and testing on another (e.g., PRONOSTIA or NASA IMS). This should be standard practice to see how well models can work in different operating conditions. Methods such as domain adaptation and transfer learning can bridge the gap between academic research and industry with real world data. The model should also provide probabilistic forecasts instead of giving exact predictions alone. Methods like Bayesian neural networks, deep ensembles, and quantile regression can provide a wide range of alternatives for predictions of RUL. This helps researchers to compare the replacement cost of the component against the risks of unexpected failures. Risk aware modeling is especially important for applications where safety is critical. Fourthly, interpretability and hybrid modeling require more consideration. Tools such as SHAP (SHapley Additive exPlanations) values, attention mechanisms, and saliency maps can clarify which features or time windows influence the RUL predictions most. Hybrid network architectures that incorporate degradation models within machine learning algorithms (as seen in [30]) offer another way toward greater clarity and estimation ability. These approaches can make prognostics not only accurate but also explainable and

trustworthy to researchers. Future research should address the current lack of uncertainty quantification and limited interpretability in bearing RUL prediction. Integrating probabilistic methods such as Monte Carlo dropout, Bayesian inference, or ensemble variance can provide confidence intervals to improve decision reliability, while applying interpretability techniques like SHAP and attention mechanisms—as demonstrated in recent studies can enhance model transparency and tackle the open challenge of feature attribution in time-series data. Finally, Industrial implementation will require lightweight devices capable of operating in real time. To make machine learning models lighter and faster without losing accuracy, we should explore techniques such as knowledge distillation, quantization, and pruning. It's also important to test these models in various real-world conditions, like dealing with noisy signals, sensor failures, or missing data, to confirm they work consistently. When evaluating these models, we should look at how they affect real-world outcomes, like improving reliability, reducing downtime, and saving on maintenance costs, in addition to just technical performance. This approach will help us move these predictive models from the lab to actual industry use, showcasing their real value.

A promising direction is to couple AI-based RUL estimation with adaptive backstepping or nonlinear optimal control frameworks, enabling machines to automatically compensate for degradation effects in real time. Such an integrated architecture would transform traditional predictive maintenance into an adaptive, self-correcting process — improving reliability, safety, and operational efficiency in industrial automation. In summary, the reviewed studies demonstrate notable advances in RUL prediction of bearings, with deep learning and hybrid approaches achieving high accuracy on standard datasets. Nevertheless, the field still faces challenges of generalization, reproducibility, interpretability, and deployment readiness. By investing in developing diverse datasets, advancing preprocessing techniques, adopting probabilistic and understandable models, and prioritizing deployment-focused research, the researchers can transform promising study into reliable, real-world prognostic systems.

6 Conclusion

This review shows how research on predicting RUL of bearings has advanced over the past two decades. The review shows that the research on RUL of Bearing using AI started with the use of conventional machine learning methods that depend on carefully extracted vibration features has now moved toward deep learning and hybrid models capable of extracting features or patterns directly from raw signals itself. This change has brought significant improvements in predictive maintenance plan with early models such as SVMs and multilayer

perceptrons generally reached only 85–90% accuracy, while more recent CNN–LSTM hybrids and BiLSTM networks regularly exceed 95%, with some reporting almost error-free performance on benchmark datasets. Approaches based on generative AI models, including GANs and autoencoders, have further improved the classification by handling challenges of noisy data and the shortage of samples from various failure conditions. Nevertheless, progress on open-source datasets does not fully resolve the challenges of applying these algorithms in real world conditions. A substantial dependance on PHM2012, PRONOSTIA, NASA IMS, and FEMTO-ST has created a testing environment that, while useful, does not reflect the complexity of real-world conditions, where load, speed, lubrication, and environmental conditions differ significantly. Many studies miss significant aspects, like the imbalance in datasets and the lack of complete examples of how equipment fails in real world scenario. The inconsistency in reporting how data is processed, along with different ways of measuring success, makes it hard to compare results across studies. There are also challenges with understanding how deep learning models work since they often act like black boxes. Plus, not enough focus is given to how uncertain the predictions are, how efficiently these models can be deployed, and the limitations of using them in real-time conditions. Overall, we see both successes and gaps. On the positive side, modern models show that artificial intelligence can greatly improve how we predict bearing failures. However, for these technologies to be used in the real world, we need to invest in diverse datasets, standardize how we process and evaluate data, create models that are easier to understand, and develop solutions that work reliably in real-time. By tackling these challenges, we can connect promising research with practical applications in industry. This review highlights how far we've come and what needs to be done to ensure that our tools for predicting bearing failures are useful, trustworthy, and scalable in real engineering environments.

Authors' contribution

Rajkumar Palaniappan is the solely responsible for the conception, searching relevant articles, writing and finalization of the manuscript

Acknowledgment

I would like to thank the University of Technology Bahrain management for providing adequate resources and time for preparing this review article.

References

- [1] R. Palaniappan, "Predictive manufacturing: State of the art, design standards, limitations,

- challenges, and future perspectives,” 2024, p. 020004. doi: 10.1063/5.0235999.
- [2] 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,” *Mech Syst Signal Process*, vol. 104, pp. 799–834, 2018, doi: <https://doi.org/10.1016/j.ymsp.2017.11.016>.
- [3] I. Nejjar, F. Geissmann, M. Zhao, C. Taal, and O. Fink, “Domain adaptation via alignment of operation profile for Remaining Useful Lifetime prediction,” *Reliab Eng Syst Saf*, vol. 242, p. 109718, 2024, doi: <https://doi.org/10.1016/j.res.2023.109718>.
- [4] Z. Zhao, Bin Liang, X. Wang, and W. Lu, “Remaining useful life prediction of aircraft engine based on degradation pattern learning,” *Reliab Eng Syst Saf*, vol. 164, pp. 74–83, 2017, doi: <https://doi.org/10.1016/j.res.2017.02.007>.
- [5] L. Guo, N. Li, F. Jia, Y. Lei, and J. Lin, “A recurrent neural network based health indicator for remaining useful life prediction of bearings,” *Neurocomputing*, vol. 240, pp. 98–109, 2017, doi: <https://doi.org/10.1016/j.neucom.2017.02.045>.
- [6] C. Ferreira and G. Gonçalves, “Remaining Useful Life prediction and challenges: A literature review on the use of Machine Learning Methods,” *J Manuf Syst*, vol. 63, pp. 550–562, 2022, doi: <https://doi.org/10.1016/j.jmsy.2022.05.010>.
- [7] L. Zhang, Z. Mu, and C. Sun, “Remaining Useful Life Prediction for Lithium-Ion Batteries Based on Exponential Model and Particle Filter,” *IEEE Access*, vol. 6, pp. 17729–17740, 2018, doi: 10.1109/ACCESS.2018.2816684.
- [8] Y. Fan, S. Nowaczyk, and T. Rögnvaldsson, “Transfer learning for remaining useful life prediction based on consensus self-organizing models,” *Reliab Eng Syst Saf*, vol. 203, p. 107098, Nov. 2020, doi: 10.1016/j.res.2020.107098.
- [9] M. Yan, X. Wang, B. Wang, M. Chang, and I. Muhammad, “Bearing remaining useful life prediction using support vector machine and hybrid degradation tracking model,” *ISA Trans*, vol. 98, pp. 471–482, Mar. 2020, doi: 10.1016/j.isatra.2019.08.058.
- [10] M. Baptista, E. M. P. Henriques, I. P. de Medeiros, J. P. Malere, C. L. Nascimento, and H. Prendinger, “Remaining useful life estimation in aeronautics: Combining data-driven and Kalman filtering,” *Reliab Eng Syst Saf*, vol. 184, pp. 228–239, Apr. 2019, doi: 10.1016/j.res.2018.01.017.
- [11] Z. Chen, Y. Li, T. Xia, and E. Pan, “Hidden Markov model with auto-correlated observations for remaining useful life prediction and optimal maintenance policy,” *Reliab Eng Syst Saf*, vol. 184, pp. 123–136, 2019, doi: <https://doi.org/10.1016/j.res.2017.09.002>.
- [12] M. Elmahallawy, T. Elfouly, A. Alouani, and A. M. Massoud, “A Comprehensive Review of Lithium-Ion Batteries Modeling, and State of Health and Remaining Useful Lifetime Prediction,” *IEEE Access*, vol. 10, pp. 119040–119070, 2022, doi: 10.1109/ACCESS.2022.3221137.
- [13] M. S. Rathore and S. P. Harsha, “Prognostic Analysis of High-Speed Cylindrical Roller Bearing Using Weibull Distribution and k - Nearest Neighbor,” *J Nondestruct Eval Diagn Progn Eng Syst*, vol. 5, no. 1, Feb. 2022, doi: 10.1115/1.4051314.
- [14] Y. Zhou, M. Huang, and M. Pecht, “Remaining useful life estimation of lithium-ion cells based on k -nearest neighbor regression with differential evolution optimization,” *J Clean Prod*, vol. 249, p. 119409, 2020, doi: <https://doi.org/10.1016/j.jclepro.2019.119409>.
- [15] R. and N. S. K. and N. N. M. and V. V. and N. F. G. Abdalla Muhamed Abdelhamed Bakhet and Palaniappan, “An Overview of Industry 5.0: Key Features, Advantages, Limitations and Future Directions,” in *Tech Fusion in Business and Society: Harnessing Big Data, IoT, and Sustainability in Business: Volume 1*, R. K. Hamdan, Ed., Cham: Springer Nature Switzerland, 2025, pp. 633–643. doi: 10.1007/978-3-031-84628-1_53.
- [16] S. Khan, T. Yairi, S. Tsutsumi, and S. Nakasuka, “A review of physics-based learning for system health management,” *Annu Rev Control*, vol. 57, p. 100932, 2024, doi: <https://doi.org/10.1016/j.arcontrol.2024.100932>.
- [17] R. Palaniappan, “Cyber-Physical Systems: Design Standards, Applications, Limitations, Challenges, and Future Perspectives,” 2025, pp. 17–30. doi: 10.1007/978-3-031-77078-4_2.
- [18] R. Huang, L. Xi, X. Li, C. Richard Liu, H. Qiu, and J. Lee, “Residual life predictions for ball bearings based on self-organizing map and back propagation neural network methods,” *Mech Syst Signal Process*, vol. 21, no. 1, pp. 193–207, Jan. 2007, doi: 10.1016/j.ymsp.2005.11.008.
- [19] S. L. Chen, M. Craig, R. Callan, H. Powrie, and R. Wood, “Use of Artificial Intelligence Methods for Advanced Bearing Health Diagnostics and Prognostics,” in *2008 IEEE Aerospace Conference*, IEEE, Mar. 2008, pp. 1–9. doi: 10.1109/AERO.2008.4526604.
- [20] C. Castejón, O. Lara, and J. C. García-Prada, “Automated diagnosis of rolling bearings using MRA and neural networks,” *Mech Syst Signal Process*, vol. 24, no. 1, pp. 289–299, Jan. 2010, doi: 10.1016/j.ymsp.2009.06.004.
- [21] C. Sun, Z. Zhang, and Z. He, “Research on bearing life prediction based on support vector machine and its application,” *J Phys Conf Ser*, vol. 305, p. 012028, Jul. 2011, doi: 10.1088/1742-6596/305/1/012028.
- [22] H.-E. Kim, A. C. C. Tan, J. Mathew, and B.-K. Choi, “Bearing fault prognosis based on health state probability estimation,” *Expert Syst Appl*, vol. 39, no. 5, pp. 5200–5213, Apr. 2012, doi: 10.1016/j.eswa.2011.11.019.

- [23] Z. Tian, “An artificial neural network method for remaining useful life prediction of equipment subject to condition monitoring,” *J Intell Manuf*, vol. 23, no. 2, pp. 227–237, Apr. 2012, doi: 10.1007/s10845-009-0356-9.
- [24] K. Medjaher, D. A. Tobon-Mejia, and N. Zerhouni, “Remaining Useful Life Estimation of Critical Components With Application to Bearings,” *IEEE Trans Reliab*, vol. 61, no. 2, pp. 292–302, Jun. 2012, doi: 10.1109/TR.2012.2194175.
- [25] X. Chen, Z. Shen, Z. He, C. Sun, and Z. Liu, “Remaining life prognostics of rolling bearing based on relative features and multivariable support vector machine,” *Proc Inst Mech Eng C J Mech Eng Sci*, vol. 227, no. 12, pp. 2849–2860, Dec. 2013, doi: 10.1177/0954406212474395.
- [26] A. Soualhi, K. Medjaher, and N. Zerhouni, “Bearing Health Monitoring Based on Hilbert–Huang Transform, Support Vector Machine, and Regression,” *IEEE Trans Instrum Meas*, vol. 64, no. 1, pp. 52–62, Jan. 2015, doi: 10.1109/TIM.2014.2330494.
- [27] J. Ben Ali, B. Chebel-Morello, L. Saidi, S. Malinowski, and F. Fnaiech, “Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network,” *Mech Syst Signal Process*, vol. 56–57, pp. 150–172, May 2015, doi: 10.1016/j.ymsp.2014.10.014.
- [28] P. Bangalore and L. B. Tjernberg, “An Artificial Neural Network Approach for Early Fault Detection of Gearbox Bearings,” *IEEE Trans Smart Grid*, vol. 6, no. 2, pp. 980–987, Mar. 2015, doi: 10.1109/TSG.2014.2386305.
- [29] M. Elforjani, “Estimation of Remaining Useful Life of Slow Speed Bearings Using Acoustic Emission Signals,” *J Nondestr Eval*, vol. 35, no. 4, p. 62, Dec. 2016, doi: 10.1007/s10921-016-0378-0.
- [30] L. Ren, J. Cui, Y. Sun, and X. Cheng, “Multi-bearing remaining useful life collaborative prediction: A deep learning approach,” *J Manuf Syst*, vol. 43, pp. 248–256, Apr. 2017, doi: 10.1016/j.jmsy.2017.02.013.
- [31] J. Deutsch and D. He, “Using Deep Learning-Based Approach to Predict Remaining Useful Life of Rotating Components,” *IEEE Trans Syst Man Cybern Syst*, vol. 48, no. 1, pp. 11–20, Jan. 2018, doi: 10.1109/TSMC.2017.2697842.
- [32] L. Ren, Y. Sun, H. Wang, and L. Zhang, “Prediction of Bearing Remaining Useful Life With Deep Convolution Neural Network,” *IEEE Access*, vol. 6, pp. 13041–13049, 2018, doi: 10.1109/ACCESS.2018.2804930.
- [33] H. O. Omoregbee, M. U. Olanipekun, and A. B. Edward, “Non Coding of Big Dataset and the use of Neural Network Regression Artificial Intelligence Model in Azure for Predicting the Remaining Useful Life (RUL) of Bearing,” in *2019 IEEE AFRICON*, IEEE, Sep. 2019, pp. 1–7. doi: 10.1109/AFRICON46755.2019.9133738.
- [34] C. Cheng *et al.*, “A Deep Learning-Based Remaining Useful Life Prediction Approach for Bearings,” *IEEE/ASME Transactions on Mechatronics*, vol. 25, no. 3, pp. 1243–1254, Jun. 2020, doi: 10.1109/TMECH.2020.2971503.
- [35] B.-L. Lu, Z.-H. Liu, H.-L. Wei, L. Chen, H. Zhang, and X.-H. Li, “A Deep Adversarial Learning Prognostics Model for Remaining Useful Life Prediction of Rolling Bearing,” *IEEE Transactions on Artificial Intelligence*, vol. 2, no. 4, pp. 329–340, Aug. 2021, doi: 10.1109/TAI.2021.3097311.
- [36] H. Ding, L. Yang, Z. Cheng, and Z. Yang, “A remaining useful life prediction method for bearing based on deep neural networks,” *Measurement*, vol. 172, p. 108878, Feb. 2021, doi: 10.1016/j.measurement.2020.108878.
- [37] Q. Wang, S. Wang, B. Wei, W. Chen, and Y. Zhang, “Weighted K-NN Classification Method of Bearings Fault Diagnosis With Multi-Dimensional Sensitive Features,” *IEEE Access*, vol. 9, pp. 45428–45440, 2021, doi: 10.1109/ACCESS.2021.3066489.
- [38] J. A. de Leon, X. Batiller, K. C. Fernandez, M. M. Vaay, C. R. Renosa, and R. Concepcion, “Remaining Useful Life Estimation for Roller Bearings through a Takagi-Sugeno Fuzzy Logic Model with Convolutional Neural Networks Consequents,” in *2022 IEEE 14th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*, IEEE, Dec. 2022, pp. 1–6. doi: 10.1109/HNICEM57413.2022.10109404.
- [39] H. S. Kumar and G. Upadhyaya, “Fault diagnosis of rolling element bearing using continuous wavelet transform and K- nearest neighbour,” *Mater Today Proc*, vol. 92, pp. 56–60, 2023, doi: <https://doi.org/10.1016/j.matpr.2023.03.618>.
- [40] M. Motahari-Nezhad and S. M. Jafari, “Comparison of MLP and RBF neural networks for bearing remaining useful life prediction based on acoustic emission,” *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, vol. 237, no. 1, pp. 129–148, Jan. 2023, doi: 10.1177/13506501221106556.
- [41] R. Palaniappan, “Comparative Analysis of Support Vector Machine, Random Forest and k-Nearest Neighbor Classifiers for Predicting Remaining Usage Life of Roller Bearings,” *Informatica*, vol. 48, no. 7, Apr. 2024, doi: 10.31449/inf.v48i7.5726.
- [42] L. Magadán, J. C. Granda, and F. J. Suárez, “Robust prediction of remaining useful lifetime of bearings using deep learning,” *Eng Appl Artif Intell*, vol. 130, p. 107690, Apr. 2024, doi: 10.1016/j.engappai.2023.107690.
- [43] Y. Ma, “Bearing Life Prediction Model for Electromechanical Equipment by Integrating Deep Neural Network and K-Nearest Neighbor

- Algorithm and Its Application,” *Journal of Theoretical and Applied Mechanics*, vol. 62, no. 4, pp. 721–735, Oct. 2024, doi: 10.15632/jtam-pl/194237.
- [44] U. Farooq, M. Ademola, and A. Shaalan, “Comparative Analysis of Machine Learning Models for Predictive Maintenance of Ball Bearing Systems,” *Electronics (Basel)*, vol. 13, no. 2, p. 438, Jan. 2024, doi: 10.3390/electronics13020438.
- [45] K. Bharatheedasan, T. Maity, L. A. Kumaraswamidhas, and M. Durairaj, “Enhanced fault diagnosis and remaining useful life prediction of rolling bearings using a hybrid multilayer perceptron and LSTM network model,” *Alexandria Engineering Journal*, vol. 115, pp. 355–369, Mar. 2025, doi: 10.1016/j.aej.2024.12.007.
- [46] V. I. Vlachou *et al.*, “Intelligent Fault Diagnosis of Ball Bearing Induction Motors for Predictive Maintenance Industrial Applications,” *Machines*, vol. 13, no. 10, p. 902, Oct. 2025, doi: 10.3390/machines13100902.
- [47] Y. Jin, D. Liu, Y. Xiao, and L. Cui, “Dual-channel dynamic spline graph convolutional network for bearing remaining useful life prediction,” *Reliab Eng Syst Saf*, vol. 266, p. 111731, Feb. 2026, doi: 10.1016/j.ress.2025.111731.