Tribology and Materials
Vol. 4, No. 2, 2025, pp. 100-115

ISSN 2812-9717
https://doi.org/10.46793/tribomat.2025.009

Performance comparison of machine learning algorithms for condition monitoring of tapered roller bearings

Harshal AHER *, Nilesh GHUGE

Matoshri College of Engineering and Research Center, Eklahare, India
*Corresponding author: harshal.aher5@gmail.com

Keywords
tapered rolling bearings
bearing defect
ANN
vibration
condition monitoring

History
Received: 14-04-2025
Revised: 19-05-2025
Accepted: 27-05-2025

Abstract

This paper investigated the implementation of machine learning algorithms for health monitoring and fault detection of tapered roller bearings (TRBs) (30205 J2/Q, 30206 J2/Q and 30207 J2/Q). Three defect models were considered: inner race defect, outer race defect and roller defect, in addition to data from the healthy bearings condition. An L27 orthogonal array design was used to generate a comprehensive dataset for each defect model, considering various operational parameters such as load, unbalance, defect type, bearing type and speed. Kurtosis was extracted as the sole feature from the vibration signals for fault classification. Several machine learning models, including artificial neural network (ANN), decision tree, support vector machine (SVM), random forest, adaptive boosting (AdaBoost), extreme gradient boosting (XGBoost), gradient boosting and categorical boosting (CatBoost), were employed to predict fault severity. The results show that the ANN model accurately predicts faults based on the kurtosis metric. This study illustrates the capability of machine learning, particularly ANN, in enhancing the predictive maintenance strategies for TRBs, thereby enabling early fault detection under varying operational conditions.

1. Introduction

Rolling element bearings play a critical role in rotating machinery but are susceptible to early failure due to both mechanical and tribological challenges. Extensive research and international standards identify several root causes, including inadequate lubrication, misalignment, excessive loading, defects in material or manufacturing defects, environmental exposure such as dust and moisture and poor installation practices. Notably, nearly 80 % of bearing failures are associated with problems related to lubrication and wear mechanisms [1]. Bearing failures can lead to significant operational challenges, increased maintenance costs and potential safety concerns. As a result, precisely and promptly identifying defects in bearings is crucial to preventing unexpected equipment failures and maintaining uninterrupted operations.

Tapered roller bearings (TRBs) are a specialised type of bearing that is extensively used across a wide range of industrial sectors. This is because they have a distinctive capacity to withstand both radial and axial loads, making them well-suited for numerous applications [2,3]. Bearing defects can take various geometric forms, including circular indentations, square pits and linear spalls across the bearing surfaces. It can experience a range of potential problems, including difficulties with its inner, outer and rolling components [3]. These defects can develop due to variables such as overloading, misalignment, lubrication issues and wear.

Conventional maintenance practices, such as scheduled inspections and timely addressing of problems, frequently fail to detect problems in the early stages. Consequently, monitoring methods that analyse vibration, acoustic signals and thermal data have gained significance for identifying bearing problems early in the series [4]. The analysis of vibration patterns is a commonly employed technique for evaluating the condition

of rotating equipment because it can assist in the early identification of potential bearing problems [5]. Various analytical methods, such as time-domain, frequency-domain and time-frequency analyses, are employed to diagnose bearing problems [6,7]. Jamadar and Vakharia [8] employed analysis of variance (ANOVA) to examine the influence of multiple factors, such as load, speed and defect geometry, on the considered vibration characteristics. Zarour et al. [7] used the ANOVA approach and Tukey's post hoc tests to identify effective indicators for detecting bearing deterioration in the early stages.

Prior research on TRBs has mainly focused on investigating lubrication, contact dynamics and misalignment. The data signals were analysed using a variety of analytical techniques, including kurtosis, root mean square, peak levels and fast Fourier transform, to obtain valuable and informative attributes from the dataset [9]. The kurtosis statistic, a measure that captures the peak and heavy-tailed distribution of the data, can offer valuable information about the bearing's condition. This is because higher kurtosis values are often linked to the presence of localised defects within the bearing components [10]. Advanced methods, such as empirical mode decomposition, dimensional analysis, continuous wavelet transform, wavelets transform and fuzzy logic, have demonstrated effectiveness in identifying and diagnosing bearing defects. Giraudo et al. [11] developed a validated multibody model for spherical roller bearings with defects. It aids in generating synthetic datasets for training fault diagnosis algorithms.

Gunerkar et al. [12] proposed a method that combines the wavelet transform with an artificial neural network (ANN) to effectively detect bearing faults, which is crucial for predictive maintenance. Their study contrasts this approach with the traditional *k*-nearest neighbors algorithm, which is a simple supervised learning technique that assigns class labels based on the majority vote of the *k* nearest data points. The findings demonstrate that the ANN-based method offers improved fault classification accuracy. Kumar et al. [13] developed an ANN model that effectively predicted the performance of a journal bearing under micropolar lubrication, thereby reducing the time required compared to traditional methods. Bankar et al. [14] studied deep learning models using vibration data to accurately identify faults in TRBs, achieving a 99 % success rate and outperforming traditional methods. The proposed deep learning approach is both efficient and effective for the real-time diagnosis of bearing failures. Narendiranath et al. [15] demonstrated that both ANN and deep neural network (DNN) deliver strong classification performance. The DNN model achieved a perfect 100 % accuracy in diagnosing faults in self-aligning bearings. Similarly, the ANN approach yielded a highly accurate 95.7 % success rate. Ali et al. [16] used the empirical mode decomposition (EMD) and ANN approach, which effectively identifies diverse bearing faults.

Leveraging run-to-failure vibration data enables real-time monitoring. Vijay et al. [17] demonstrated that wavelet-based diagnosing substantially enhances the signal-to-noise ratio, resulting in more effective feature extraction and more precise bearing fault classification. The methods proved successful with both synthetic and experimental data, and the ANN and support vector machine (SVM) models attained high classification accuracy by utilising denoised signals. Ertunc et al. [18] conclude that the combined application of ANN and flexible neuro-fuzzy analytical systems offers a powerful and reliable approach for bearing fault diagnosis. The adaptive neuro-fuzzy inference system model's enhanced capacity to manage uncertainty and complex patterns delivers superior performance in fault detection compared to ANN methods alone.

Kaplan et al. [19] findings demonstrated that the ANN-based model is capable of achieving a perfect 100 % accuracy in classifying the size of bearing faults, even when working with real-time data. Deák et al. [20] demonstrated that the Symlet-5 and Morlet wavelet techniques were successful in identifying faults and estimating their sizes. Using the proposed approach, the maximum error in estimating defect size was a modest 4.12 %. Vibration data collected from the test setup was leveraged to validate the technique, with the findings further corroborated through optical microscopy and direct measurements. Tiwari et al. [21] optimised the bearing design, which improves both durability and dynamic capabilities, showcasing the effectiveness of the genetic algorithm approach in creating high-quality, reliable bearings. Sensitivity analysis verifies that the optimal design parameters do not negatively impact the bearing's dynamic performance.

Önel and Benbouzid [22] presented a methodology that effectively identifies bearing failures through experimental assessments

involving deliberately introduced bearing damage. The radial basis function ANN model simplifies the fault detection and diagnosis process, demonstrating impressive accuracy in recognising bearing defects. Kankar et al. [23] found that ANN and SVM are effective tools for automated bearing fault diagnosis. Additionally, the research revealed that bearings with rough inner race surfaces exhibit chaotic vibrations, providing useful insights for detecting severe defects. Narendiranath Babu et al. [24] highlighted the superior effectiveness of DNN compared to conventional ANN in this field. DNN can deliver improved results by incorporating additional layers, ultimately enhancing the accuracy of fault classification. Patil et al. [25] demonstrate that integrating discrete wavelet transform and ANN yields highly precise bearing fault identification, facilitating efficient real-time condition monitoring and proactive maintenance.

Samanta and Al-Balushi [26] used the ANN approach, which has emerged as an effective solution for accurately identifying bearing faults, thereby facilitating efficient real-time condition monitoring. Kumar et al. [27] proposed a technique utilising Shannon entropy and continuous wavelet transform, which proves to be an effective way of identifying defects and estimating their size in TRBs. Samanta et al. [28] examine and compare the ability of various neural network models, such as multi-layer perceptron, radial basis functions and probabilistic neural networks, to accurately detect bearing faults using vibration data. Additionally, the researchers used genetic algorithms to optimise the configurations of these neural networks, which, in turn, enhanced the overall accuracy of the fault detection process. The investigators used carefully controlled methods, including electrical discharge machining (EDM) and laser technology, to create artificial defects. This enables a consistent and comprehensive investigation of how bearing vibration patterns vary under different defect conditions [29].

In the present study, a detailed investigation was conducted on three TRBs (30205 J2/Q, 30206 J2/Q and 30207 J2/Q). Failures were most frequently noticed in the inner and outer raceways and rolling elements, primarily due to surface fatigue, pitting, corrosion and fretting wear [30]. This research addresses these issues by using precise mounting methods combined with vibration-based condition monitoring through kurtosis analysis, supported by both industry standards and experimental validation. The main objective is to predict kurtosis values using several machine learning algorithms, including ANN, decision tree, SVM, random forest, adaptive boosting (AdaBoost), extreme gradient boosting (XGBoost), gradient boosting and categorical boosting (CatBoost). The parameters for this study were selected based on previous research emphasising their significance in the vibration behaviour of TRBs. The applied algorithm examines the effects of unbalance, load, speed and defects (circular, square and line defects) on vibration responses [31,32].

2. Experimentation setup and data acquisition

2.1 Experimentation setup

The experimental apparatus employed in this study, as shown in Figure 1, was created by Tyrannus Innovative Engineering and Research Academy [33]. It consists of a shaft driven by a three-phase induction motor, two supports, a flexible shaft coupling, bearings and a variable-speed drive controller to regulate the rotational speed. The coupling serves to accommodate and compensate for any potential misalignment between the driven shaft and the motor output shaft. A healthy bearing is mounted at the motor side, while the sample bearing is positioned at the opposite end. A triaxial accelerometer is mounted on the support, as shown in Figure 1, to capture the vibration data of the sample bearing.

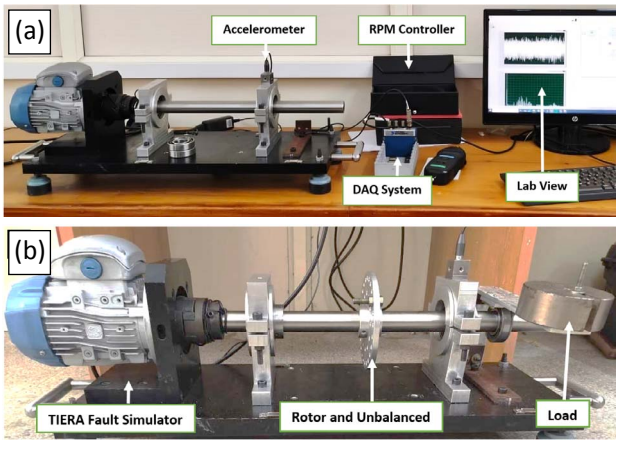

Figure 1. Test setup: (a) healthy condition and (b) condition with artificial defect and unbalanced

A dedicated data acquisition card NI-9234 (National Instruments, USA) is used to interface with accelerometer sound and vibration measurement. This card is housed in the cDAQ-

Table 1. Specification of test bearings

Bearing parameter	Designation		
	30205 J2/Q	30206 J2/Q	30207 J2/Q
Inner race diameter, mm	25	30	35
Outer race diameter, mm	52	62	72
Total width, mm	16.25	17.25	18.25
Inner ring width, mm	15	16	17
Outer ring width, mm	13	14	15
Net weight, kg	0.149	0.228	0.326
Number of rollers	17	17	17
Contact angle, °	14.036	14.036	14.036
Pitch diameter, mm	38.5	46.0	53.5
Roller diameter, mm	6.34	7.75	8.90
Roller length, mm	10	11	12

9178 chassis (National Instruments, USA), which is linked to a LabVIEW-enabled computer using a USB cable. A LabVIEW function has been implemented to capture the analogue acceleration signal in both the time and frequency domains and to store the data for subsequent analysis. The test bearings examined in this study were TRBs with designations of 30205 J2/Q, 30206 J2/Q and 30207 J2/Q. The specifications of these test bearings are presented in Table 1.

Localised defects of circular, square and line geometries with specified dimensions shown in Table 2 and outlined in Figures 2a, 2b and 2c were artificially induced on the inner and outer races and on the roller of the test bearings using CNC-controlled Sparkonix ZNC EDM machine as illustrated in Figure 2d. Defects were created with a discharge current of 6 A, a 220 V voltage, a 100 µs pulse on-time and a 50 µs pulse off-time. A low-viscosity, highly refined mineral-based dielectric fluid was used to ensure adequate cooling. A copper-tungsten electrode was employed for consistent and repeatable defect creation.

Table 2. Dimensions of bearing defects

No.	Defect position	Defect size, mm	Defect geometry
1	inner race	1 × through (diameter × depth)	circular
2		1.5 × 1.5 × through (width × length × depth)	square
3		2 × 7 × through (width × length × depth)	line
4	outer race	1 × 1 (diameter × depth)	circular
5		1.5 × 1.5 × 1.5 (width × length × depth)	square
6		2 × 7 × 2 (width × length × depth)	line
7	roller	1 × 1 (diameter × depth)	circular
8		1.5 × 1.5 × 1.5 (width × length × depth)	square
9		2 × 7 × 2 (width × length × depth)	line

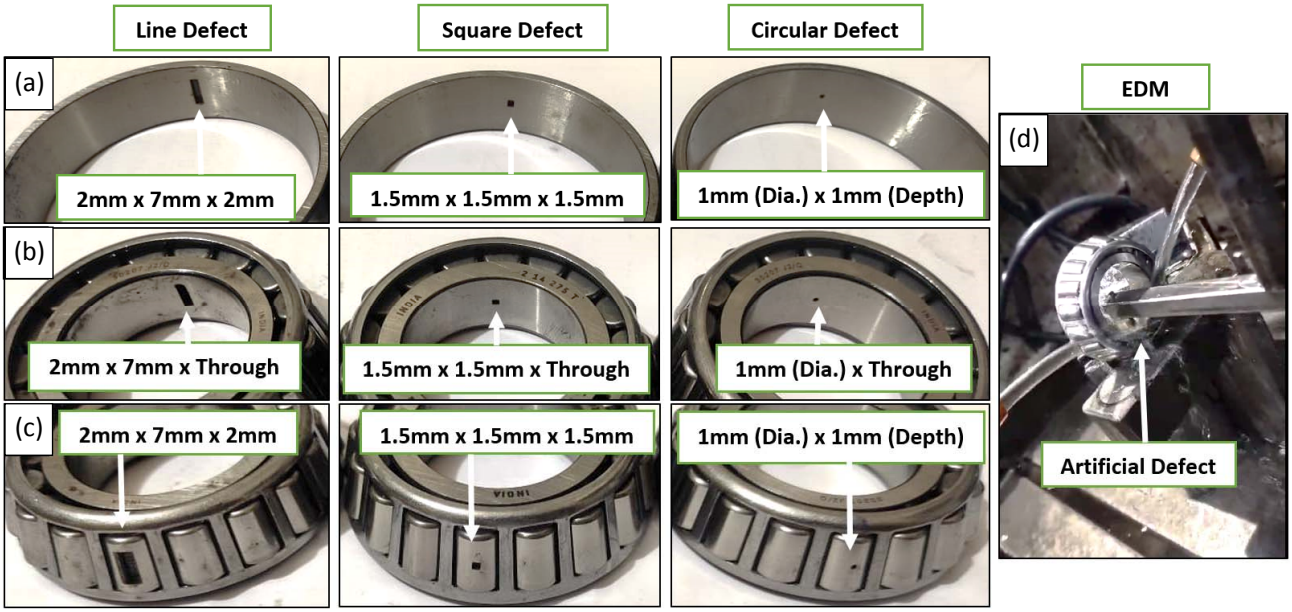

Figure 2. Artificially induced defects: a) defects on outer race, b) defects on inner race, c) defects on roller and d) defects creation using EDM

Table 3. Input variables in bearing test

Parameter	Value
Bearing number	30205 | 30206 | 30207
Load, kg	1 | 2 | 3
Speed, rpm	1000 | 1500 | 2000
Unbalance, g	0 | 10 | 15 | 20
Defect geometry	circular | square | line | no defect
Defect position	inner race | outer race | roller

To investigate the vibration behaviour of the test bearings under varying operating conditions (Table 3), the radial loads were adjusted to 1, 2 and 3 kg. Similarly, the shaft rotational speed was controlled using a variable frequency drive, enabling speeds of 1000, 1500 and 2000 rpm. Additionally, unbalanced conditions were created by attaching concentrated masses of 10, 15 and 20 g to the rotor. The test bearings, both with and without artificially induced defects, were reinstalled individually in the experimental setup to enable the collection of vibration data.

To enhance experimental efficiency, the Taguchi L27 orthogonal array was employed, reducing the need for a full factorial design by limiting the number of experiments to just 27 trials for each defect category, including inner race defect (IRD), outer race defect (ORD) and roller defect (RD). This design facilitated a structured examination of five independent variables (bearing properties, rotational speed, applied load,

Table 4. Input and output parameters as per design of experiments (DOE) and experimental results for inner race defect (IRD), outer race defect (ORD) and roller defect (RD) conditions

Bearing number	Load, kg	Speed, rpm	Defect geometry	Unbalance, g	IRD kurtosis	ORD kurtosis	RD kurtosis
30205	1	1000	circular	10	3.3585	3.4200	3.2236
				15	3.4952	3.7900	3.2756
				20	3.4542	3.5400	3.8115
	2	1500	square	10	3.2878	3.6140	3.5479
				15	3.4906	3.7038	3.5022
				20	2.9809	3.6260	4.3186
	3	2000	line	10	3.3771	5.2200	5.4082
				15	3.5868	5.3700	5.2111
				20	4.1201	5.3500	5.8958
30206	1	1500	line	10	3.5030	4.7400	3.6807
				15	3.9145	4.8712	4.0279
				20	3.6986	5.0730	4.4125
	2	2000	circular	10	3.4880	4.4800	3.6689
				15	3.4615	4.9227	3.8911
				20	3.1051	5.1348	4.1023
	3	1000	square	10	2.9003	3.1766	4.2060
				15	3.0428	3.1163	4.4892
				20	2.9493	3.2422	4.5132
30207	1	2000	square	10	7.6461	4.6320	4.5417
				15	7.3319	4.8124	4.4783
				20	7.3841	4.6041	5.2232
	2	1000	line	10	6.3073	3.9266	4.0451
				15	6.9017	4.4387	4.2305
				20	7.1030	4.2641	4.7295
	3	1500	circular	10	6.1085	3.0565	4.8800
				15	5.9853	3.5798	4.9105
				20	6.3131	3.6229	5.5335

Table 5. Kurtosis value of a healthy bearing condition

Bearing number	Load, kg	Speed, rpm	Kurtosis
30205	1	1000	2.1973
	2	1500	2.3294
	3	2000	2.8951
30206	1	1500	2.4908
	2	2000	2.7908
	3	1000	2.2998
30207	1	2000	2.8830
	2	1000	2.4520
	3	1500	2.6654

unbalance and defect shape), each at three levels, thereby considerably reducing the experimental work while maintaining statistical validity. This resulted in a dataset of 81 samples, along with 9 samples from healthy conditions. The kurtosis values were then extracted for these test cases and used as input features for the machine-learning models (Tables 4 and 5).

2.2 Data acquisition

Time-domain analysis is recognised as the superior and elementary method for examining vibration signals. Numerous time-domain features, such as kurtosis, crest factor, root mean square and peak-to-peak value, could be leveraged for health monitoring and defect detection [16]. Kurtosis is a time-domain statistical feature that measures the sharpness or peakedness of a vibration signal. It helps in identifying impulsive faults in rotating machinery [34]. In this study, the kurtosis value K is calculated from the acquired vibration data using MATLAB. This behaviour can be determined using Equation (1).

$$\mathrm{K} = \frac{\sum_{i=1}^{N}(y_i - \bar{y})^4}{N\sigma^4}, \tag{1}$$

where K is kurtosis, y_i is instantaneous magnitude, $\bar{y}$ is mean of vibration signals, N is sample length and σ is standard deviation [35].

Empirical research has established the high diagnostic utility of kurtosis, as it displays a marked increase in the presence of bearing defects compared to healthy bearing conditions. Various regression approaches were employed to predict the bearing's vibration kurtosis. The input variables are presented in Table 3, and the output variable is presented in Table 6. The vibration kurtosis was selected as the target variable, with the defect geometry serving as one of the input factors for the analysis. Furthermore, four standard key performance indicators were used to measure the model's effectiveness, i.e. mean squared error (MSE), mean absolute error (MAE), mean absolute percentage error (MAPE) and the coefficient of determination (R^2). These performance indicators are defined by the following equations.

$$\mathrm{MSE} = \frac{1}{m}\sum_{i=1}^{m}(P_i - T_i)^2, \tag{2}$$

$$\mathrm{MAE} = \frac{1}{m}\sum_{i=1}^{m}\left|P_i - T_i\right|, \tag{3}$$

$$\mathrm{MAPE} = \frac{1}{m}\sum_{i=1}^{m}\left|\frac{P_i - T_i}{T_i}\right|, \tag{4}$$

$$\mathrm{R}^2 = \frac{\sum_{i=1}^{m}(P_i - T_i)^2}{\sum_{i=1}^{m}(P_i - \bar{T})^2}, \tag{5}$$

where P_i indicates the predicted value, T_i is the actual observed value, $\bar{T}$ denotes the mean of the actual values and m refers to the total number of data points.

Table 6. Output variables in bearing test

Parameter	Min	Max	Mean	SD
Vibration kurtosis	2.1973	7.6461	3.91104	1.26237

SD – standard deviation

Equation (2) defines MSE, which calculates the average of squared differences between predicted and actual values, thereby emphasising larger errors. Equation (3) presents MAE, which averages the absolute deviations without considering direction. MAPE, shown in Equation (4), evaluates the mean percentage difference between predicted and actual values, offering a relative error scale. Equation (5) describes R^2, which measures how well the predicted values account for the variation in the actual data. Values of R^2 closer to 1 suggest better model performance. These evaluation metrics are widely adopted for assessing model prediction accuracy across regression-based applications [36].

3. ANN modelling

Various supervised machine learning techniques are based on statistical learning theory.

They have proven to be valuable in categorisation and regression tasks, particularly in scenarios with limited sample sizes, like those encountered in fault diagnosis applications [23]. The SVM is a type of intelligent computing method similar to ANN that can perform similar tasks. However, SVM employs a distinct approach. Instead of a complex neural network structure, the SVM points a hyperplane across the two types of data, effectively dividing the data associated with respective types. The ANN attempts to reduce the error in the training dataset, a process known as empirical risk minimisation. In contrast, the SVM approach is based on risk regularisation and focuses on minimising the upper bound of the transferability error [18].

Decision tree is a widely used machine-learning technique. It offers excellent transparency, flexibility and the ability to select relevant features. Unlike many other complex machine-learning models that can be difficult to understand, decision tree generates a user-friendly, tree-like structure that is easy for people to interpret. The ability to interpret the model's decision-making process is particularly valuable in industries such as healthcare, law and investment, where transparency is important.

Another supervised machine learning algorithm is the random forest, which builds upon the decision tree approach with enhancements. Random forest is a powerful algorithm that helps overcome the problem of overfitting, which can occur with decision tree models. Overfitting occurs when a model performs with high accuracy on training data but performs poorly on unseen data. The random forest approach addresses this by maintaining low bias and low variance, making it a robust and reliable choice for predictive tasks [37].

The AdaBoost algorithm is a widely used machine learning technique. It works by addressing the errors made by weak classifiers, and it is generally less prone to overfitting compared to many other learning algorithms. The recognition performance of AdaBoost-based classifiers is typically promising [38]. Boosting is a machine learning approach that can be applied to both regression and categorisation tasks. It works by building a series of simple models, with each new model aiming to address the limitations of the previous ones. When these individual models are built based on the gradient of the loss function, this technique is referred to as a gradient boosting algorithm.

The XGBoost algorithm, a powerful machine learning system for tree boosting, was employed. There are three key reasons why XGBoost outperforms the other three boosting methods: First, it uses a regularised loss function, which helps prevent overfitting. Second, it can scale down the weights of individual trees, reducing the impact of any single tree on the ultimate prediction. Third, it utilises column sampling, a technique similar to the random forest approach, which further enhances the model's performance [39].

CatBoost is an enhanced gradient-boosted decision tree framework comparable to XGBoost. This algorithm effectively addresses the issues of gradient bias and prediction shift by automatically treating discrete features as numerical characteristics. It leverages a combination of features that exploit the relationships among them, significantly expanding the feature dimensions. In addition, CatBoost employs a symmetrical tree model design to mitigate overfitting and improve the accuracy and generalisability of the algorithm [40].

The ANN model was implemented using the Keras Sequential API in TensorFlow. Other models, such as SVR, decision tree, random forest, AdaBoost and gradient boosting, were developed using Scikit-learn, while XGBoost and CatBoost were implemented using their respective libraries.

The ANN model encompasses a multistage process, including gathering the data, preprocessing it, partitioning the data into training and testing sets and ultimately applying the ANN training and testing algorithms to analyse the proposed model [41]. Figure 3 illustrates the flowchart of the ANN regression analysis process. The neurons in the network accept multiple inputs and produce a single output. To represent the mathematical operations occurring within these neurons, the weighted sum of the inputs can be expressed using Equation (6). The ANN features interconnected neurons, resembling the intricate structure of the human nervous system. The neurons in the network accept multiple inputs and produce a single output [42]. To represent the mathematical operations occurring within these neurons n, the weighted sum of the inputs.

$$n = \sum_{i=1}^{p} (W_i X_i + b), \tag{6}$$

where p expresses the number of inputs, W_i denotes the input weights, X_i denotes the inputs and b denotes the bias value [43].

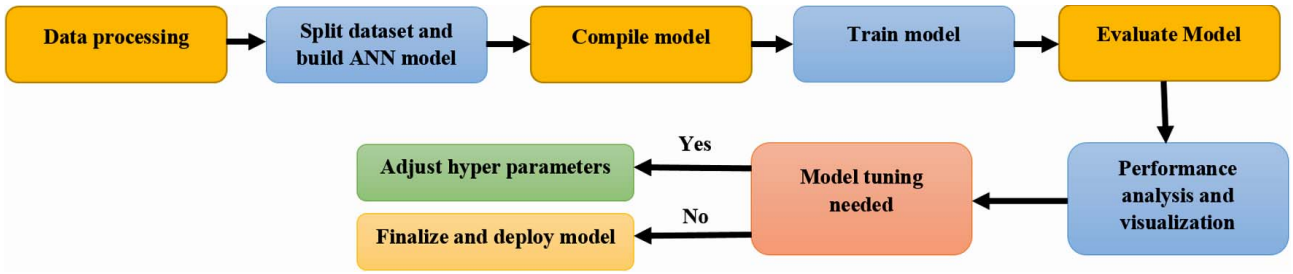

Figure 3. ANN regression analysis process flowchart

The weighted inputs and biased sum of the neuron values are passed through an activation function. This results in the neuron's output. When the nonlinear function is denoted as $f(n)$, the equation can be expressed as:

$$f(n) = f(\sum_{i=1}^{p}(W_i X_i + b)). \quad (7)$$

The categorical data were converted into a numerical format using one-hot encoding. The encoder transformed the three categorical columns "defect geometry", "defect position" and "bearing number" into numerical one-hot encoded columns. Visualising the data using a box plot can help identify outliers and understand the distribution of the "vibration kurtosis" variable. The seaborn library was utilised to create a box plot to highlight any outliers present in this column.

The analysis identified several outliers in the "vibration_kurtosis" column using the interquartile range method. These outliers may suggest unusual vibration patterns in the data. The outliers in the "vibration_kurtosis" column were addressed by capping them at the lower and upper bounds. In addition, the standard scalar was utilised to normalise the features in X by eliminating the mean and scaling to unit variance. Standardising the features by removing the mean and scaling to unit variance helps prevent the models from being overly influenced by features with large values. This normalisation confirms that all features have a mean of 0 and a standard deviation of 1, which improves the overall model capabilities. The data was partitioned into 80 % for training and 20 % for testing, as outlined by a test size variable set to 0.2. This setup allows the model to learn from the training data and then be assessed on new, unseen data. Fixing the random state was set to 42 to ensure that the data split remains consistent across runs, promoting the reproducibility of the results. This data partition strategy helps assess the model's effectiveness on new, previously unknown data before deployment.

The Keras sequential method is a way to build neural networks by stacking layers arranged vertically in a linear alignment. This model type is often used as a starting point for creating multi-layer neural network architectures. The neural network method featured a total of five layers: three hidden layers with 64, 32 and 16 neurons, respectively, an output layer with a single neuron and an input layer defined by the input features. This fully connected architecture allowed the method to learn complicated patterns from the data efficiently. The initial layer takes 14 features, as indicated by the input shape (None, 14), where "None" represents a batch size. The first hidden layer is configured with 64 neurons and employs the ReLU activation function, which brings in nonlinear capabilities. The output shape (None, 64) signifies that each data sample is transformed into a 64 dimensional vector at this stage. The following hidden layer had 32 neurons, and the layer after that contained 16 neurons. This gradual decrease in the number of neurons allowed the model to effectively learn and capture important patterns from the data. The output layer has a single neuron, as the model was designed for a regression task aimed at predicting a continuous target variable. The Adam optimisation algorithm, an advanced technique, was employed to dynamically adjust the learning rates during training. Figure 4 illustrates the design of an ANN that employs the Keras sequential model. It offers a visual representation of how data moves through the different layers of the network. The neural network had four fully linked layers, called dense layers. The number of neurons decreased in every subsequent layer (64 → 32 → 16 → 1), allowing the model to effectively study important patterns from the dataset. The output layer consisted of a single neuron that was used to predict the kurtosis value.

4. Results

The ANN model employed the Adam optimiser, an advanced technique that dynamically adjusted

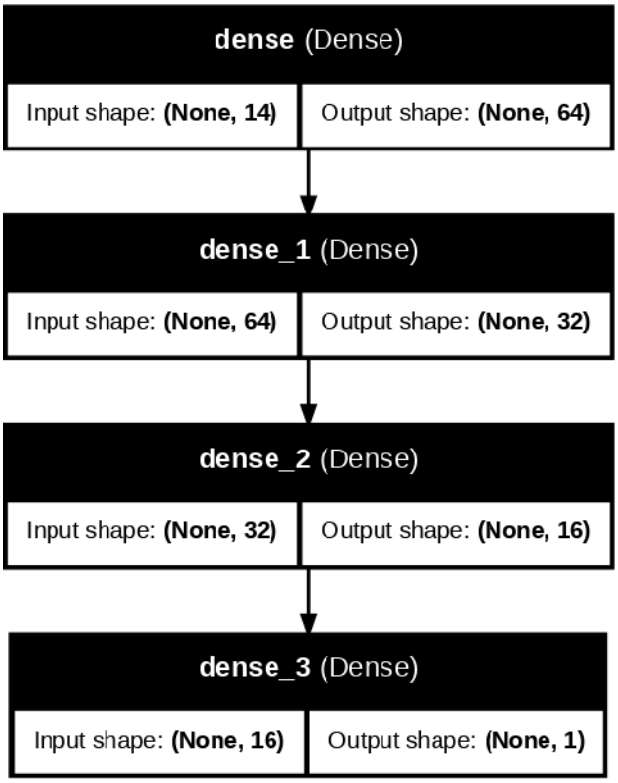

Figure 4. Configuration of ANN model

the learning rates for each parameter during the training procedure. This helped reduce the loss function by efficiently updating the model's weights. Adam is an advanced optimisation algorithm that modifies the learning rate for each variable based on the first-order (mean) and second-order (variance) statistics of past gradients. This improves the model's stability and makes it well-suited for deep learning tasks. The model used MSE as the error function, which quantified the mean of the squared deviation between the actual and predicted values, making it well-suited for regression tasks. In addition, the team monitored MAE as a performance measure to assess the model's predictive accuracy, measuring the average absolute deviation between the predicted and actual values.

The model was trained on the training data, with each training round going through the entire dataset. This iterative training process was repeated 50 times. The model updated its weights after analysing small batches of 5 samples at a time. Furthermore, the team evaluated the model's effectiveness using a separate verification dataset, which provided insights into its ability to simplify new, unknown data. The training history, which recorded details such as loss, validation loss and other relevant metrics, was leveraged to monitor the model's learning progression and detect any signs of overfitting. Figure 5 illustrates the training loss and MAE across 50 epochs for an ANN regression model. The vertical axis displays the values of the error function and the MAE metric. At the start, the model's performance was poor, as evidenced by the high initial values. The horizontal axis shows the number of training epochs, with the model being trained over 50 epochs in total. The loss starts at a high value but then falls rapidly in the first 10 epochs. The graph shows the model's performance improving as training progresses. The loss and MAE metrics both decrease rapidly at first and then stabilise at low values, indicating that the model is learning effectively and making increasingly accurate predictions. This suggests the model trained well overall.

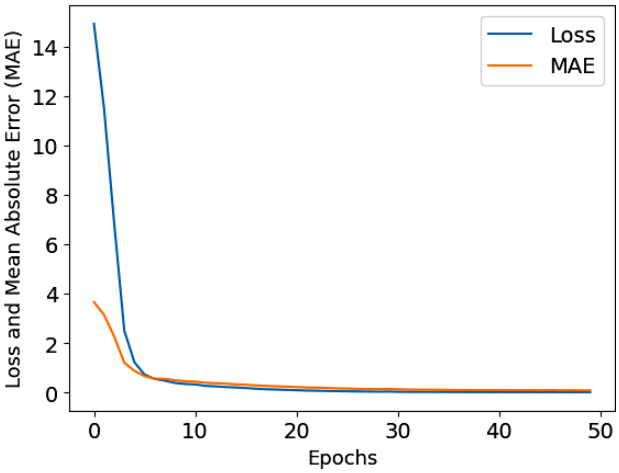

Figure 5. Training loss and MAE plot vs. epochs

The effectiveness of the trained ANN regression method was assessed on test data. The MAE measures the average absolute deviation between predicted and actual values. A lower value of MAE indicates that the model's predictions are more accurate. The obtained MAE of 0.1341 suggests that the model makes relatively low errors on average. The MSE estimates the standard deviation of prediction errors. A smaller MSE means the model's predictions are more accurate. The MSE of 0.0410 is relatively small, implying the model's prediction errors are very low. The R^2 value shows how effectively the method describes the variance in the data. R^2 ranges from 0 to 1, with greater values reflecting stronger model performance. The R^2 of 0.9701 shows that the model accounts for 97.01 % of the variation in the test data, which is an excellent fit. Figure 6 compares the actual and the predicted values for vibration kurtosis in a regression model. The blue line depicts the actual measured values from the dataset and the orange

line shows the predicted values generated by the trained ANN. The *x*-axis displays the sample index, representing distinct test instances selected from the 20 % hold-out dataset used to test the ANN model, while the *y*-axis shows the kurtosis value. Because the predicted values are much closer to the actual values, the model is accurate and reliable. The MAE of 0.1341 and the R^2 value of 0.9701 both indicate the model's strong predictive capability.

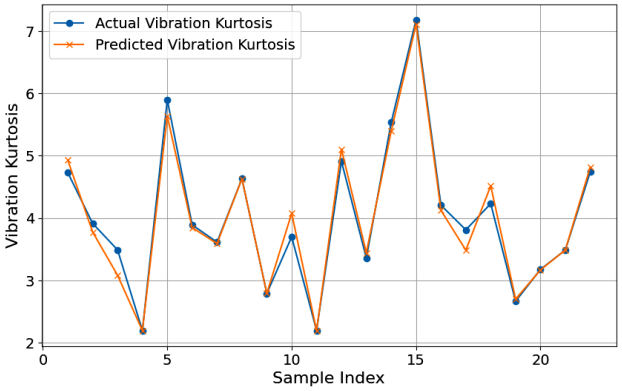

Figure 6. Actual vs. predicted vibration kurtosis for different test conditions

The trained ANN model was employed to forecast the vibration kurtosis for a given input, which represented the scaled sensor data from a bearing. The input was preprocessed in the same manner as the training data. After obtaining the predicted vibration kurtosis, the code categorised the bearing conditions into five groups, as shown in the following table. The ANN model classifies the bearing vibration [26], which is continuously measured by an accelerometer, into distinct fault types based on the predictions made by the smart models. The values and their associated magnitudes in the desired output matrix are given in Table 7.

5. Evaluating and validating the machine learning models

The effectiveness of the ANN model has been compared to other machine learning algorithms (Table 8). The results indicate that the ANN model outperforms other methods, such as random forest, decision tree, SVM, AdaBoost, XGBoost, gradient boosting and CatBoost, in accurately predicting vibration kurtosis. This suggests that the ANN model is the most effective in correctly identifying faulty bearings based on the obtained data. In addition, several well-known machine learning algorithms are compared, as shown in Figure 7.

Table 7. Categories of bearing defect conditions

Category	Condition
Normal condition	predicted kurtosis < 3
Defect is initiated	3 ≤ predicted kurtosis < 4
Defect is moderate	4 ≤ predicted kurtosis < 6
Defect is significant	6 ≤ predicted kurtosis < 10
Severe fault	predicted kurtosis ≥ 10

The evaluated methods include decision tree, SVM, random forest, AdaBoost, XGBoost, gradient boosting and CatBoost. Among these, a decision tree is a single learning algorithm [40], while random forest, AdaBoost, XGBoost, gradient boosting and CatBoost are combined learning techniques. However, SVM is a supervised learning algorithm that does not fit single or ensemble learning categories. The optimal model configuration was identified through a grid search and a 3-fold cross-validation process (Table 9). The effectiveness measures, including MAE, MSE and R^2, were then calculated. To provide a clearer illustration, the predicted values are plotted against the experimental values in Figure 7. The $y = x$ line in the plot represents the ideal situation in which the predicted and tested values are perfectly aligned. The graph shows a clear linear relationship between the predicted and tested values, with only a small amount of scatter or variation noticed. The majority of the data points align closely with the perfect fit line, indicating that the ANN model makes highly accurate predictions.

Table 8. Performance measures for bearing kurtosis prediction

Measure		ANN	Decision tree	SVM	Random forest	AdaBoost	XGBoost	Gradient boosting	CatBoost
R^2	training	0.9934	0.82132	0.99861	0.98544	0.92234	0.99995	1	0.99005
	testing	0.9701	0.73175	0.95941	0.89715	0.86260	0.94960	0.89827	0.94346
MSE	training	0.0102	0.27920	0.00217	0.02275	0.12135	0.00008	0	0.01555
	testing	0.0410	0.36700	0.05553	0.14071	0.18797	0.06893	0.13919	0.07736
MAE	training	0.0697	0.39924	0.04501	0.10644	0.29891	0.00545	0.00101	0.09951
	testing	0.1341	0.44139	0.18448	0.25617	0.32619	0.18914	0.24940	0.19090

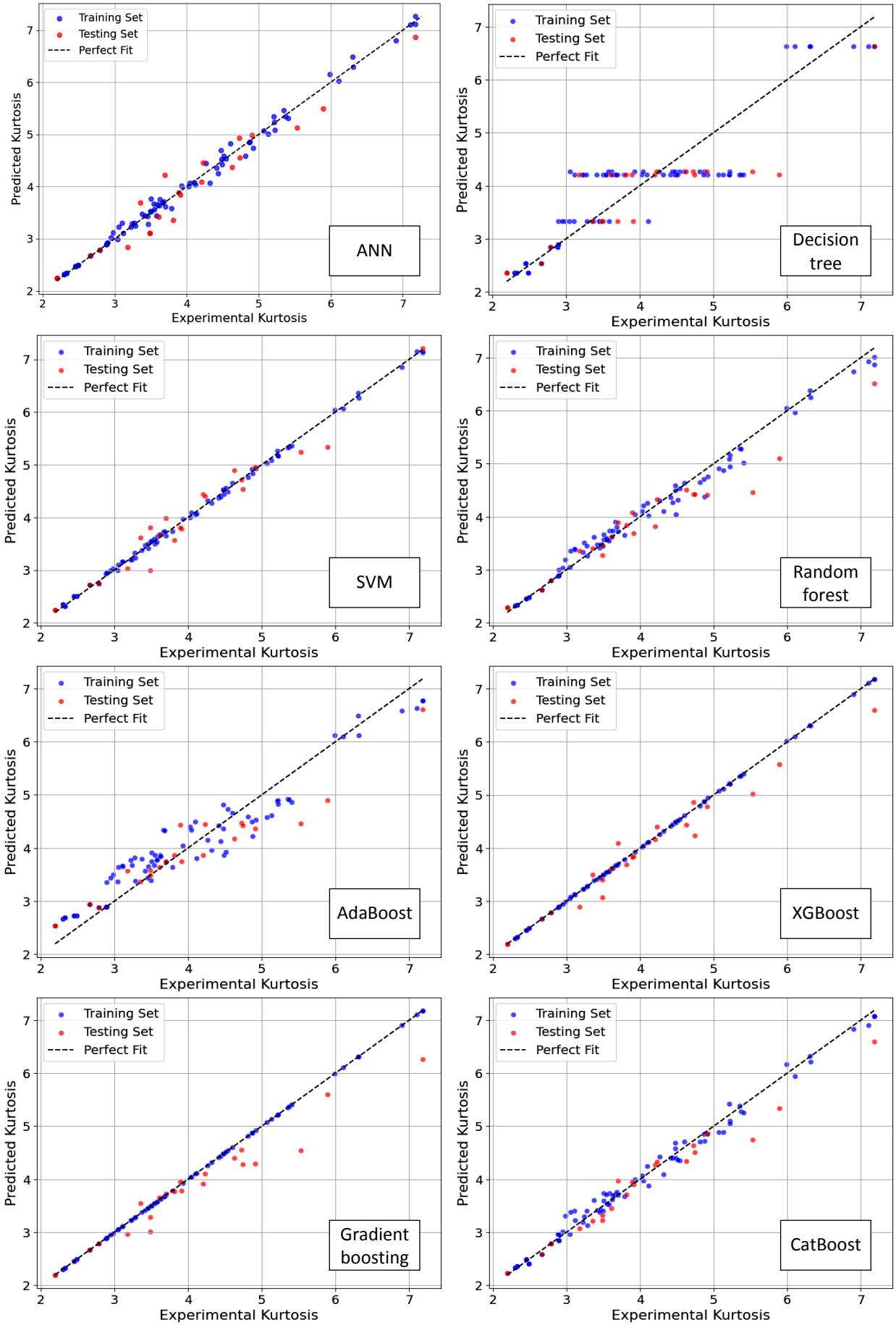

Figure 7. Comparing the predicted values and the corresponding experimental measurements for the different machine learning models

Table 9. Model parameter settings used for the various algorithms

Category	Model	Parameter	Value
Single	decision tree	maximum depth of tree	3
		minimum samples for split	4
		minimum samples of leaf node	3
		cross-validation	3
	SVM	C	1000
		kernel	RBF
Ensemble	random forest	number of estimators	250
		maximum depth of tree	30
		minimum samples for split	2
		minimum samples of leaf node	1
		cross-validation	3-fold
	AdaBoost	number of estimators	450
		learning rate	0.45
		maximum depth of the tree	3
		minimum samples for split	2
		minimum samples of leaf node	1
		cross-validation	3-fold
	XGBoost	number of estimators	450
		learning rate	0.12
		maximum depth of the tree	11
		minimum samples for split	1
		cross-validation	3-fold
	gradient boosting	number of estimators	30
		learning rate	0.2
		maximum depth of the tree	5
		minimum samples for split	5
		cross-validation	3
	CatBoost	number of iterations	200
		learning rate	0.1
		maximum depth of the tree	3
		cross-validation	3

The quantitative measures presented in Table 8 demonstrate the model's strong performance. For the training set, the model achieved impressive scores, with an R^2 value of 0.9934, an MSE value of 0.0102 and an MAE value of 0.0697. Similarly, the testing set results were also excellent, with an R^2 value of 0.9701, an MSE of 0.0410 and an MAE value of 0.1341. These metrics demonstrate that the model possesses both strong learning capabilities and dependable predictive performance. The results indicate that the ANN model, a single learning algorithm, demonstrates superior performance compared to ensemble learning techniques in predicting the vibration kurtosis of bearings. As shown in Figure 7 and Table 8, the ANN model achieved the highest R^2 value of 0.9701, the lowest MSE of 0.0410 and the lowest MAE of 0.1341 for the testing set, outperforming other machine learning algorithms such as random forest, decision tree, SVM, AdaBoost, XGBoost, gradient boosting and CatBoost. The ANN achieves the best balance between training and testing performance with high R^2 (0.9701), low MSE (0.0410) and low MAE (0.1341). To validate the observed performance difference, a 10-fold cross-validation followed by statistical testing was conducted, where both the paired t-test ($p = 0.0133$) and the Wilcoxon signed-rank test ($p = 0.0059$) indicated that the ANN model significantly outperforms XGBoost in terms of R^2 accuracy. The gradient boosting algorithm performs perfectly on the training data ($R^2 = 1$, MSE = 0), but the testing metrics were weaker. The decision tree has the lowest R^2, the highest MSE and the highest MAE on the testing data.

6. Discussion

The effectiveness of different machine learning methods, such as ANN, random forest, decision tree, SVM, AdaBoost, gradient boosting, XGBoost and CatBoost, was evaluated using the key metrics of R^2, MSE and MAE for both the training and testing datasets. The ANN model demonstrated outstanding performance, achieving an R^2 of 0.9934 in training and 0.9701 in testing, indicating a strong predictive capability that is consistent with findings reported in previous research, as illustrated in Figure 7. While gradient boosting achieved a perfect R^2 of 1 in training, its low testing accuracy of 0.89827 suggests the potential for overfitting. Overfitting can be improved by considering regularisation techniques, such as

limiting tree depth, lowering the learning rate, dropout or using early stopping. In addition, SVM, XGBoost, and CatBoost exhibited good performance, although their testing errors were slightly higher than those of the ANN model. In contrast, the decision tree model displayed the weakest performance, with R^2 values of 0.82132 in training and 0.73175 in testing, indicating limited predictive ability. The ANN model demonstrated good performance on the training and testing datasets, indicating minimal signs of overfitting. In contrast, while XGBoost and gradient boosting achieved near-perfect scores on the training data, their testing performance exhibited a more noticeable decline, suggesting the potential for overfitting. Meanwhile, random forest and AdaBoost showed moderate performance. However, their higher MSE and MAE on the testing set suggest opportunities for further improvement.

The findings from the ANN model in this study, highlighted by an R^2 value of 0.9701 and an MAE of 0.1341 on the test dataset, indicate a high level of predictive accuracy. These results align well with previous research, where advanced machine learning methods have achieved accuracies as high as 99 % or even 100 % in bearing fault detection [14,15]. While earlier studies mainly concentrated on classification tasks, this work demonstrates that comparable performance can be attained in regression-based condition monitoring, specifically when estimating kurtosis values. Other research incorporating empirical mode decomposition with ANN has also successfully identified bearing faults, although it did not report regression metrics such as R^2 or MAE [16]. Moreover, although the model ensemble algorithms such as XGBoost and CatBoost, which are less commonly evaluated in earlier studies that primarily used traditional methods, have shown promising results in previous research. Their results here were slightly inferior to that of the ANN model [41,44]. This suggests the superior generalisation ability of the ANN in this application. Overall, these findings corroborate and extend existing knowledge by confirming that ANN offers precise and dependable fault diagnosis for TRBs under varying operational conditions. The vertical bar plot (Figure 8) reveals a comparison of the performance of various machine learning models based on the R^2 score, which indicates the percentage of variance explained by each model. ANN and XGBoost show the highest explanatory power among the evaluated models.

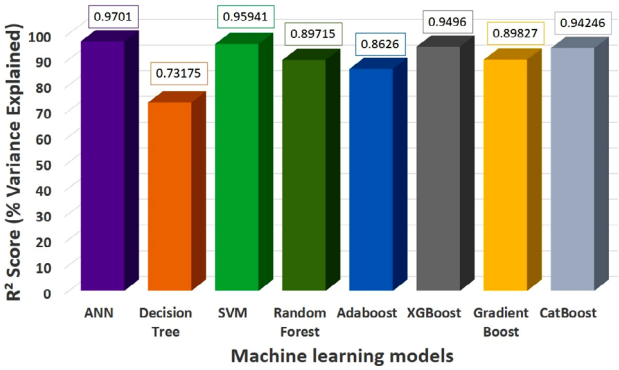

Figure 8. Comparison of performance of machine learning models

7. Conclusion

The ANN model demonstrated strong predictive capability, as evidenced by the close agreement between its predictions and actual values. The concentration of data points near the ideal fit line indicates both high accuracy and effective generalisation. Minor variations in the test data suggest that further improvements can be made by applying techniques like regularisation or expanding the training dataset.

This study evaluated eight prominent machine learning algorithms: ANN, decision tree, SVM, random forest, AdaBoost, XGBoost, gradient boosting and CatBoost. While tree-based models offer interpretability, XGBoost is particularly noted for its robustness and regularisation features. Ensemble methods such as AdaBoost and CatBoost enhance predictive performance by combining multiple models. Among all these approaches, the ANN consistently delivered superior accuracy and generalisation in forecasting vibration kurtosis.

Future research should emphasise strategies to minimise overfitting, including dropout, early stopping and cross-validation, while further optimising the ANN and increasing the diversity of training data. Exploring more advanced network architectures, such as deep neural networks, convolutional neural networks or transformer-based models, may provide additional performance gains. Furthermore, incorporating a broader range of features beyond kurtosis and applying rigorous statistical validation can improve the robustness of diagnostic models.

Acknowledgement

The authors would like to express their sincere gratitude to Matoshri College of Engineering and Research Center, Eklahare, Nashik (Affiliated to

Savitribai Phule Pune University, Pune) for providing the necessary resources and support for this research. We also extend our appreciation to our colleagues and faculty members for their valuable insights and constructive feedback.

References

[1] A. Vencl, V. Gašić, B. Stojanović, Fault tree analysis of most common rolling bearing tribological failures, IOP Conference Series: Materials Science and Engineering, Vol. 174, 2017, Paper 012048, DOI: 10.1088/1757-899X/174/1/012048

[2] L. Yang, T. Xu, H. Xu, Y. Wu, Mechanical behavior of double-row tapered roller bearing under combined external loads and angular misalignment, International Journal of Mechanical Sciences, Vol. 142-143, 2018, pp. 561-574, DOI: 10.1016/j.ijmecsci.2018.04.056

[3] A.V. Dube, L.S. Dhamande, P.G. Kulkarni, Vibration based condition assessment of rolling element bearings with localized defects, International Journal of Scientific & Technology Research, Vol. 2, No. 4, 2013, pp. 149-155.

[4] H.P. Mishra, A. Jalan, Analysis of faults in rotor-bearing system using three-level full factorial design and response surface methodology, Noise & Vibration Worldwide, Vol. 52, No. 11, 2021, pp. 365-376, DOI: 10.1177/09574565211030711

[5] L.S. Dhamande, M.B. Chaudhari, Bearing fault diagnosis based on statistical feature extraction in time and frequency domain and neural network, International Journal of Vehicle Structures and Systems, Vol. 8, No. 4, 2016, pp. 229-240, DOI: 10.4273/ijvss.8.4.09

[6] P.K. Kankar, S.C. Sharma, S.P. Harsha, Vibration signature analysis of a high speed rotor supported on ball bearings due to localized defects, Journal of Vibration and Control, Vol. 19, No. 12, 2013, pp. 1833-1853, DOI: 10.1177/1077546312448506

[7] D. Zarour, S. Meziani, M. Thomas, Experimental studies for bearings degradation monitoring at an early stage using analysis of variance, Diagnostyka, Vol. 19, No. 4, 2018, pp. 81-87, DOI: 10.29354/diag/94985

[8] I.M. Jamadar, D.P. Vakharia, An in-situ synthesized model for detection of defective roller in rolling bearings, Engineering Science and Technology, an International Journal, Vol. 19, No. 3, 2016, pp. 1488-1496, DOI: 10.1016/j.jestch.2016.05.003

[9] C. Grover, N. Turk, Rolling element bearing fault diagnosis using empirical mode decomposition and Hjorth parameters, Procedia Computer Science, Vol. 167, 2020, pp. 1484-1494, DOI: 10.1016/j.procs.2020.03.359

[10] Y. Wang, J. Xiang, R. Markert, M. Liang, Spectral kurtosis for fault detection, diagnosis and prognostics of rotating machines: A review with applications, Mechanical Systems and Signal Processing, Vol. 66-67, 2016, pp. 679-698, DOI: 10.1016/j.ymssp.2015.04.039

[11] L. Giraudo, L.G. Di Maggio, L. Giorio, C. Delprete, Dynamic multibody modeling of spherical roller bearings with localized defects for large-scale rotating machinery, Sensors, Vol. 25, No. 8, 2025, Paper 2419, DOI: 10.3390/s25082419

[12] R.S. Gunerkar, A.K. Jalan, S.U. Belgamwar, Fault diagnosis of rolling element bearing based on artificial neural network, Journal of Mechanical Science and Technology, Vol. 33, No. 2, 2019, pp. 505-511, DOI: 10.1007/s12206-019-0103-x

[13] S. Kumar, V. Kumar, A.K. Singh, Prediction of maximum pressure of journal bearing using ANN with multiple input parameters, Australian Journal of Mechanical Engineering, Vol. 20, No. 4, 2022, pp. 1069-1078, DOI: 10.1080/14484846.2020.1769540

[14] M.S. Bankar, B.B. Ahuja, M.R. Dhanvijay, Fault diagnostics on vibration data of taper roller bearing using deep learning algorithms, in Proceedings of the 2nd Indian International Conference on Industrial Engineering and Operations Management, 16-18.08.2022, Warangal, India, pp. 1590-1602, DOI: 10.46254/IN02.20220443

[15] B.T. Narendiranath, A. Aravind, A. Rakesh, M. Jahzan , P.D. Rama, Application of EMD ANN and DNN for self-aligning bearing fault diagnosis, Archives of Acoustics, Vol. 43, No. 2, 2018, pp. 163-175, DOI: 10.24425/122364

[16] J.B. Ali, N. Fnaiech, L. Saidi, B. Chebel-Morello, F. Fnaiech, Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals, Applied Acoustics, Vol. 89, 2015, pp. 16-27, DOI: 10.1016/j.apacoust.2014.08.016

[17] G.S. Vijay, H.S. Kumar, P. Srinivasa Pai, N.S. Sriram, B.K.N.R. Raj, Evaluation of effectiveness of wavelet based denoising schemes using ANN and SVM for bearing condition classification, Computational Intelligence and Neuroscience, Vol. 2012, No. 1, 2012, Paper 582453, DOI: 10.1155/2012/582453

[18] H.M. Ertunc, H. Ocak, C. Aliustaoglu, ANN- and ANFIS-based multi-staged decision algorithm for the detection and diagnosis of bearing faults, Neural Computing and Applications, Vol. 22, No. 1 supplement, 2013, pp. 435-446, DOI: 10.1007/s00521-012-0912-7

[19] K. Kaplan, S. Bayram, M. Kuncan, H.M. Ertunç, Feature extraction of ball bearings in time-space

and estimation of fault size with method of ANN, in Proceedings of the 16th Mechatronika 2014, 03-05.12.2014, Brno, Czechia, 295-300.

[20] K. Deák, T. Mankovits, I. Kocsis, Optimal wavelet selection for the size estimation of manufacturing defects of tapered roller bearings with vibration measurement using Shannon entropy criteria, Strojniški vestnik – Journal of Mechanical Engineering, Vol. 63, No. 1, 2017, pp. 3-14, DOI: 10.5545/sv-jme.2016.3989

[21] R. Tiwari, K.K. Sunil, R.S. Reddy, An optimal design methodology of tapered roller bearings using genetic algorithms, International Journal for Computational Methods in Engineering Science and Mechanics, Vol. 13, No. 2, 2012, pp. 108-127, DOI: 10.1080/15502287.2011.654375

[22] I.Y. Önel, M.E.H. Benbouzid, Induction motors bearing failures detection and diagnosis using a RBF ANN park pattern based method, International Review of Electrical Engineering, Vol. 3, No. 1, 2008, pp. 159-165.

[23] P.K. Kankar, Satish C. Sharma, S.P. Harsha, Fault diagnosis of ball bearings using machine learning methods, Expert Systems with Applications, Vol. 38, No. 3, 2011, pp. 1876-1886, DOI: 10.1016/j.eswa.2010.07.119

[24] T. Narendiranath Babu, A. Aravind, A. Rakesh, M. Jahzan, D. Rama Prabha, M. Ramalinga Viswanathan, Automatic fault classification for journal bearings using ANN and DNN, Archives of Acoustics, Vol. 43, No. 4, 2018, pp. 727-738, DOI: 10.24425/aoa.2018.125166

[25] A.B. Patil, J.A. Gaikwad, J.V. Kulkarni, Bearing fault diagnosis using discrete wavelet transform and artificial neural network, in Proceedings of the 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), 21-23.07.2016, Bangalore, India, pp. 399-405, DOI: 10.1109/ICATCCT.2016.7912031

[26] B. Samanta, K.R. Al-Balushi, Artificial neural network based fault diagnostics of rolling element bearings using time-domain features, Mechanical Systems and Signal Processing, Vol. 17, No. 2, 2003, pp. 317-328, DOI: 10.1006/mssp.2001.1462

[27] R. Kumar, J. Singh, S. Sharma, C. Li, G. Królczyk, E.M.T. Eldin, S. Wojciechowski, Identification of localized defects and fault size estimation of taper roller bearing (NBC_30205) with signal processing using the Shannon entropy method in MATLAB for automobile industries applications, Heliyon, Vol. 8, No. 12, 2022, Paper e12053, DOI: 10.1016/j.heliyon.2022.e12053

[28] B. Samanta, K.R. Al-Balushi, S.A. Al-Araimi, Artificial neural networks and genetic algorithm for bearing fault detection, Soft Computing, Vol. 10, No. 3, 2006, pp. 264-271, DOI: 10.1007/s00500-005-0481-0

[29] X.-w. Deng, P. Yang, J.-s. Ren, Y.-w. Yang, Rolling bearings time and frequency domain fault diagnosis method based on kurtosis analysis, in Proceedings of the 2014 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), 07-10.12.2014, Hong Kong, China, DOI: 10.1109/APPEEC.2014.7066018

[30] M. Vasić, B. Stojanović, M. Blagojević, Failure analysis of idler roller bearings in belt conveyors, Engineering Failure Analysis, Vol. 117, 2020, Paper 104898, DOI: 10.1016/j.engfailanal.2020.104898

[31] A.K. Ansari, P. Kumar, Vibration and acoustics analyses of tapered roller bearing, Journal of Vibration Engineering & Technologies, Vol. 12, No. 2, 2024, pp. 2467-2484, DOI: 10.1007/s42417-023-00991-9

[32] A.K. Ansari, P. Kumar, Vibro-acoustic analysis of defective taper roller bearings, Tribology International, Vol. 199, 2024, Paper 110044, DOI: 10.1016/j.triboint.2024.110044

[33] P.K. Samal, K. Sunil, I.M. Jamadar, R. Srinidhi, AI-enhanced fault diagnosis in rolling element bearings: A comprehensive vibration analysis approach, FME Transactions, Vol. 52, No. 3, 2024, pp. 450-460, DOI: 10.5937/fme2403450S

[34] V.C.M.N. Leite, J.G. Borges da Silva, G.F. Cintra Veloso, L.E. Borges da Silva, G. Lambert-Torres, E.L. Bonaldi, L.E. de Lacerda de Oliveira, Detection of localized bearing faults in induction machines by spectral kurtosis and envelope analysis of stator current, IEEE Transactions on Industrial Electronics, Vol. 62, No. 3, 2015, pp. 1855-1865, DOI: 10.1109/TIE.2014.2345330

[35] A.K. Ansari, P. Kumar, V. Rastogi, Dynamic behavior of aluminium matrix composite rotor on bearing vibration, Materials Today: Proceedings, Vol. 62, No. 1, 2022, pp. 233-238, DOI: 10.1016/j.matpr.2022.03.026

[36] D.-C. Feng, Z.-T. Liu, X.-D. Wang, Z.-M. Jiang, S.-X. Liang, Failure mode classification and bearing capacity prediction for reinforced concrete columns based on ensemble machine learning algorithm, Advanced Engineering Informatics, Vol. 45, 2020, Paper 101126, DOI: 10.1016/j.aei.2020.101126

[37] S. Chandrvanshi, S. Sharma, M.P. Singh, R. Singh, Bearing fault diagnosis using machine learning models, in D.K. Sharma, S.-L. Peng, R. Sharma, G. Jeon (Eds.), Micro-Electronics and Telecommunication Engineering, Springer, Singapore, 2024, pp. 219-233, DOI: 10.1007/978-981-99-9562-2_18

[38] W. Hu, W. Hu, S. Maybank, AdaBoost-based algorithm for network intrusion detection, IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Vol. 38, No. 2, 2008, pp. 577-583, DOI: 10.1109/TSMCB.2007.914695

[39] B. Pan, Application of XGBoost algorithm in hourly PM2.5 concentration prediction, IOP Conference Series: Earth and Environmental Science, Vol. 113, 2018, Paper 012127, DOI: 10.1088/1755-1315/113/1/012127

[40] M. Luo, Y. Wang, Y. Xie, L. Zhou, J. Qiao, S. Qiu, Y. Sun, Combination of feature selection and CatBoost for prediction: The first application to the estimation of aboveground biomass, Forests, Vol. 12, No. 2, 2021, Paper 216, DOI: 10.3390/f12020216

[41] T. Mariprasath, K.R. Cheepati, M. Rivera, Practical Guide to Machine Learning, NLP, and Generative AI: Libraries, Algorithms, and Applications, River Publishers, New York, 2024, DOI: 10.1201/9781003563945

[42] B. Samanta, K.R. Al-Balushi, S.A. Al-Araimi, Bearing fault detection using artificial neural networks and genetic algorithm, EURASIP Journal on Advances in Signal Processing, Vol. 2004, 2004, Paper 785672, DOI: 10.1155/S1110865704310085

[43] S.G. Kumbhar, R.G. Desavale, N.V. Dharwadkar, Fault size diagnosis of rolling element bearing using artificial neural network and dimension theory, Neural Computing and Applications, Vol. 33, No. 23, 2021, pp. 16079-16093, DOI: 10.1007/s00521-021-06228-8

[44] B.G. Mejbel, S.A. Sarow, M.T. Al-Sharify, L.A. Al-Haddad, A.A.F. Ogaili, Z.T. Al-Sharify, A data fusion analysis and random forest learning for enhanced control and failure diagnosis in rotating machinery, Journal of Failure Analysis and Prevention, Vol. 24, No. 6, 2024, pp. 2979-2989, DOI: 10.1007/s11668-024-02075-6