International Journal of Electrical and Computer Engineering (IJECE)
Vol. 16, No. 3, June 2026, pp. 1466~1473
ISSN: 2088-8708, DOI: 10.11591/ijece.v16i3.pp1466-1473

Bearing fault classification using decision trees and neural networks

Raid Houssem Eddine Sellaoui, Brahim Boulebtateche, Salah Bensaoula
Laboratory of Automation and Signals Annaba, Department of Electronics, Faculty of Technology, Badji Mokhtar Annaba University, Annaba, Algeria

Article Info

Article history:

Received Jun 30, 2025
Revised Feb 21, 2026
Accepted Mar 16, 2026

Keywords:

Decision trees
Diagnosis
Electric vehicle
Fault classification
Machine learning
Neural networks
Rolling bearing

ABSTRACT

In this study, we test three machine learning methodologies − binary tree, k-nearest neighbors (k-NN), and neural networks (NN) − using a range of hyperparameters. These methods are applied to a dataset consisting of extracted time series characteristics (root mean square (RMS), skewness, and kurtosis from vibration signals of various bearings subjected to different fault conditions from the intelligent maintenance systems (IMS) dataset. We evaluate how effectively these methods classify the condition of the bearings using the provided dataset. We observe the top two methods, artificial neural network (ANN) 99.29% and binary tree 98.84%. With a difference of 0.45%, the binary tree is preferred over the complex ANN due to its ease of interpretation, transparency, and minimal computation requirements. Its integration as code in embedded controllers or electronic control units (ECUs) is more efficient, which makes them faster for real-time processing and safety-critical electric vehicle (EV) systems.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Raid Houssem Eddine Sellaoui
Laboratory of Automation and Signals Annaba (LASA), Department of Electronics, Faculty of Technology, Badji Mokhtar Annaba University
Annaba, Algeria
Email: raid-houssem-eddine.sellaoui@univ-annaba.dz

1. INTRODUCTION

Advancements in rotating machinery have led to more refinement and complexity, which has raised the demand for reliability. The harsh and extreme working conditions made its parts prone to degradation. Even small defects while running a system can cause parts to break or the whole system to stop working, which can lead to major accidents and damages. Thus, rapid and precise evaluation and estimation of health status are essential for maintaining its safe and dependable operation [1]. The dependable operation of rolling bearings is vital for the accuracy and longevity of rotating machinery. Given their significance, ensuring their reliability directly impacts the overall performance of the system. Bearing degradation accounts for 40% to 50% of all failures; making it the most common type. Bearing degradation occurs at four locations: the bearing cage, rolling elements, outer race, and inner race. Second is stator failure, which happens about 30% to 35% of the time. About 10% of the time, rotors fail, making them the third most common type of failure. The last few percent are made up of all the other types of failures [2]–[4].

When defects begin to develop in a motor, they manifest as abnormalities in its operational signals (acceleration, pressure, or strain characteristics) [5]. Detecting these abnormalities is the key for early diagnosis. Among these approaches, vibration signal analysis is especially suitable for detecting faults and defects in bearings [6], [7]. The bearings produce vibration signals that display non-stationarity and periodic impact features. As a result, signal processing methods are frequently employed in engineering to recognize and extract defect characteristics from raw vibration signals for monitoring and diagnosing their condition.

Fault detection and isolation (FDI) is a branch of control engineering that focuses on upholding a system's reliability and safety during its operational period, avoiding possible breakdowns and improving overall efficiency [8]. As technology advances, defect diagnosis methods are continually developing. Modern approaches do these tasks by combining advanced signal processing and feature extraction with machine learning (ML) algorithms, such as using a Kalman filter integrated with an artificial neural network to enhance detection accuracy and enable robust fault-tolerant control [8]–[11]. A random forest classifier is another way to find out the health of a gearbox. The data are the raw vibration signals recorded from a gearbox experiment simulating six kinds of gear failures [12]. Rauber *et al.* [13] aimed to identify health conditions of rotating machines using a labeled dataset from Case Western Reserve University (CWRU). Su *et al.* [14] did a multi-fault diagnosis via support vector machine (SVM). Two approaches that achieve high diagnostic accuracy while reducing computational complexity were proposed by Alonso-González *et al.* [15] and Chen *et al.* [16]. The approach is designed to identify bearing defects with limited characteristics and observational data.

For the purposes of our work, the intelligent maintenance systems (IMS) bearing dataset [17] represented a fitting training dataset for our work. Various research studies in fields have utilized the IMS database. Various ML models have been developed to identify abnormal conditions in bearings before they lead to catastrophic failures. One study proposed a combination of convolutional neural networks (CNNs) and gated recurrent units (GRUs) for anomaly detection in rotating machinery [18]. Another study utilized a parallel long-short term memory (PARA-LSTM) model to detect anomalies in bearing vibration [19]. The data set was also used in remaining useful life (RUL) predictions using bidirectional long short-term memory (BiLSTM) neural networks [20]. It was also applied for performance evaluation in a self-selective regression model to select the most suitable to predict the RUL of the bearing [21]. Building on these foundations, this work aims for a diagnosis method using time series features extracted from the raw vibration signals. Meanwhile, the simple and efficient binary tree, k-nearest neighbors (k-NN), and artificial neural network (ANN) are employed as classifiers to shorten the training time, keeping high accuracy while lowering algorithm complexity.

2. THEORETICAL BASIS

Machine learning is employed to instruct machines in the more efficient management of data. At times, analyzing the data to discern patterns or extract information can prove challenging. We use machine learning in that case. The growing number of datasets has heightened the demand for machine learning in numerous sectors, including healthcare and defense, utilizing it to derive valuable information [22], [23].

2.1. Binary tree

A binary tree is a basic hierarchical data structure that organizes data in a tree-like format, where each node can maintain a maximum of two children. It allows efficient representation of data, which is useful in computer science applications, including but not limited to search algorithms, sorting operations, and hierarchical data modeling [24]. The binary tree's structure is characterized by a root node at the top, which serves as the entry point to the tree. Each node within the tree can contain a value or key, and the relationships among the nodes are such that the value of the left branch is typically less than that of the parent node, while the value of the right branch is greater than or equal to the parent node.

In the framework of applying binary trees for various computational tasks, specific hyperparameters can be defined to optimize the tree's construction and operation. In this study, we focus on two hyperparameters: Maximum number of splits defines the upper limit on how many times the data can be split during the construction of the binary tree. Limiting the number of splits helps in preventing overfitting, the split criterion method for determining how to split the data at each node. The three methods used in this paper are Gini's diversity index, signifying the probability for a random instance being misclassified when chosen randomly, as seen in (1); maximum deviance reduction (cross entropy), which is a measure of uncertainty or disorder, as seen in (2); and the towing rule, which is not a purity measure for a node but rather serves as an alternative criterion for determining how to split a node, as seen in (3).

$$Gini = 1 - \sum_{i=1}^{j} P(i)^2 \quad (1)$$

$$1 - \sum_{i=1}^{j} P(i).\log_2\big(P(i)\big) \quad (2)$$

$$P(L)P(R)\Big(\sum_{i=1}^{j} |L(i) - R(i)|\Big)^2 \quad (3)$$

2.2. Ensemble learning

Ensemble learning refers to the process of amalgamating multiple individual learners into a singular learner. The individual learner may include Naïve Bayes, a decision tree, or a neural network, among others, employed to reduce bias and variation. Boosting generates a strong learner by combining a series of weak learners. If a classifier shows little correlation with the correct classification, it is considered a weak learner; on the other hand, a strong learner is strongly linked with accurate classification [23]. Bagging, or bootstrap aggregating, enhances the stability and accuracy of an ML algorithm. It is applicable for regression as well as classification. Additionally, bagging helps to reduce variance and mitigate overfitting [23].

2.3. K-nearest neighbor

The k-nearest neighbor (k-NN) classification algorithm operates on the principle of proximity in multi-dimensional feature space. The main goal of k-NN classification is to give a data point a class label based on the class labels of its nearest neighbors. Specifically, the output of the algorithm is determined through a plurality vote among the object's k-nearest neighbors, where the object is assigned to the most frequently represented class of its neighbors. This non-parametric method is particularly advantageous in scenarios where the distribution of data is unknown or complex [25].

We focused on two hyperparameters to improve the performance of the k-NN classifier we used. The number of neighbors (k) specifies how many of the nearest points are utilized to classify each point during prediction, a low value of k results in a fine classifier, while a high value leads to a coarse classifier. The distance metric is the algorithm used to calculate distances between points. In this paper, we focus on two distance measures: Euclidean distance, the most widely used, is a measure of the true straight-line distance between two points in Euclidean space, as seen in (4), and Manhattan distance, also known as city block distance, calculates the sum of the absolute differences between the coordinates of two points (5) [26].

$$f(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad (4)$$

$$f(x, y) = \sum_{i=1}^{n}|x_i - y_i| \quad (5)$$

2.4. Artificial neural network

Artificial neural networks (ANNs) are simplified artificial models that are based on the biological mechanisms of neurons in human brains. An ANN is made up of several highly interconnected artificial processing units (nodes), which are grouped into layers, creating a network. The number of nodes in the input and output layers is determined by the specific number of input and output variables needed to define the issue. A trial-and-error approach typically determines the number and node density of the hidden layers [27].

Each node (except in the input layer) is calculated as the sum $\sum_{i=1}^{n}$ of its weighted inputs W_iX_i and bias θ_i. This sum is then processed by a non-linear activation function $f(NET)$ to produce the output [28]. This study uses three activation functions: Sigmoid, a common non-linear function that outputs values between 0 and 1. As seen in (7). Tanh (hyperbolic tangent), which is similar to sigmoid but is symmetric around zero, is often preferred for its gradient behavior. As seen in (9) and Figure 1. Rectified linear unit (ReLU), a non-linear function where only a subset of neurons activates at once. It is computationally efficient, though it can sometimes cause neurons to "die" (stop updating) during training [29].

$$OUT = f(NET) = f(\sum_{i=1}^{n} W_iX_i + \theta_i) \quad (6)$$

$$f(x) = \frac{1}{1+e^{-x}} \quad (7)$$

$$f(x) = \frac{2}{sigmoid(2x)} - 1 \quad (9)$$

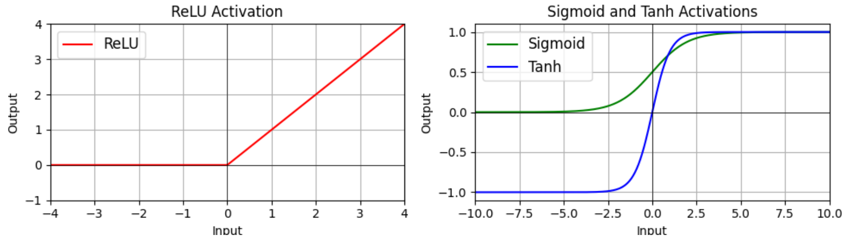

Figure 1. ReLU (Red) Sigmoid (Green) and Tanh (Blue) or logistic activation function graph

2.5. Dataset description

The intelligent maintenance systems (IMS) bearing dataset came from the test rig shown in Figure 2. It has four Rexnord ZA-2115 double-row bearings (characteristics are seen in Table 1) mounted on a shaft. An AC motor connected to the shaft by rubber belts keeps the speed of rotation at a steady 2000 rpm. A spring mechanism puts a radial load of 6,000 lbs. on the shaft and bearings. Also, a circulation system that controls both the temperature and flow lubricates all of the bearings. Each housing has two high-sensitivity quartz ICP accelerometers (PCB 353B33) that measure the x and y axes. We note that all failures happened after 100 million revolutions, which is the expected lifespan of the bearing [17].

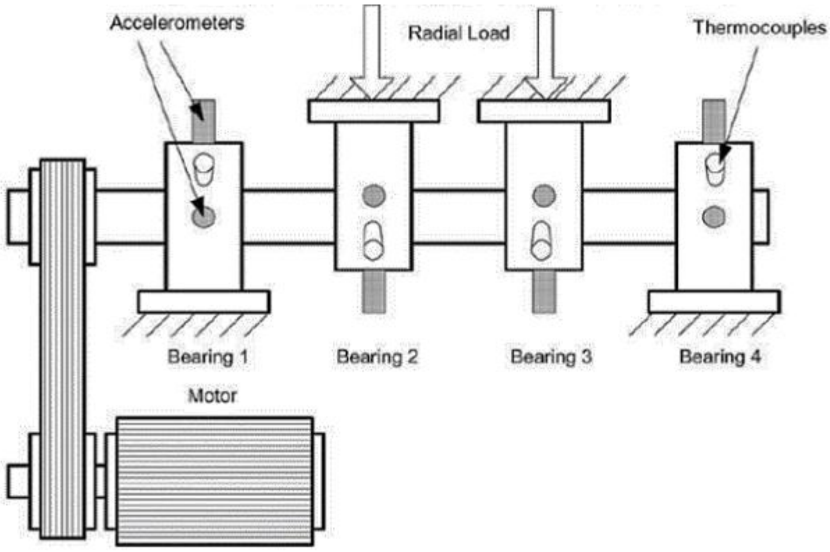

Figure 2. The test rig of IMS bearing dataset

Table 1. Characteristics of the bearings used

Rexnord ZA-2115 Characteristics		
Number of rolling element per row	16	16
Rolling element diameter	0.331 inch	8.4 mm
Pitch diameter	2.815 inch	71.5 mm
Contact angle	15.17°	15.17°
Static load	6000 lbs.	26690 N

The IMS packet contains three test-to-failure experiment datasets. Each consists of multiple files, with each containing 1-second snapshots of vibration signals recorded at set intervals. The data packet characteristics are summarized in Table 2. Each file contains 20,480 data points; at a sampling frequency of 20 kHz, data acquisition occurs at the frequency of ten minutes (the first dataset's initial forty-three files were collected every five minutes) [17].

Table 2. Datasets characteristics and summary

Dataset	Number of channels	Number of files	Endurance duration	Announced damages
Dataset1	8	2156	34 days 12h	Bearing 4: rolling element | Bearing 3: inner race
Dataset2	4	984	6 days 20h	Bearing 1: outer race
Dataset3	4	4448	31 days 10h	Bearing 3: outer race

2.6. Features for time series analysis

In the analysis of time series data, various statistical features are employed to offer information about the distribution, variability, and overall characteristics of the data (root mean square (RMS), standard deviation (STD), absolute mean (AM), kurtosis, skewness, and peak to peak) [30], all of which are used as inputs. In Table 3, we outline said key features along with their mathematical formulas.

Table 3. Time Series features and their mathematical formula

Time series features	Mathematical formulas
Root mean square (RMS)	$x_{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i^2)}$
Standard deviation (STD)	$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}$
Absolute mean (AM)	$\bar{x} = \frac{1}{N}\sum_{\{i=1\}}^{\{N\}}\lvert x_i\rvert$
Kurtosis	$K = \frac{1}{N}\sum_{i=1}^{N}\frac{(x_i - \bar{x})^4}{\sigma^4}$
Skewness	$Sk = \frac{1}{N}\sum_{i=1}^{N}\frac{(x_i - \bar{x})^3}{\sigma^3}$
Peak to peak	$x_{p2p} = \max(x) - \min(x)$

3. METHOD

In our work, all modules were designed and trained using MATLAB on a dual-core central processing unit (CPU) with a 16-GB RAM system. This setup smoothed experimentation with various model architectures and parameters, ensuring vigorous assessment of each classification method's performance.

For the decision trees, we trained several models using three split creation methods: Gini's diversity index, maximum deviance reduction, and the twoing rule. Each method was evaluated across a range of splits, varying from 4 to 500. Additionally, ensemble tree methods were assessed, including bagging, RUS-Boost, and AdaBoost. These ensemble techniques were tested with varying parameters (maximum number of splits, number of estimators (learners), and the learning rate) to determine optimal configuration for improved performance.

The k-NN classifier's modules used different distance metrics and the number of neighbors. Specifically, two distance measures: City block (Manhattan) distance and Euclidean distance. While the number of neighbors ranged from 1 to 100, allowing us to observe how the neighborhood size influences classification accuracy.

For neural network classifiers, multiple configurations were used, focusing on activation functions and network architecture. We tested three activation functions—ReLU, Tanh, and Sigmoid—applied uniformly across hidden layers. The hidden layers varied from one to three, with each having 10, 20, 50, 70, or 100 neurons per layer. Various layer arrangements were constructed, including increasing sequences (*e.g.*, 10-50-100), decreasing sequences (*e.g.*, 50-20-10), uniform layers (*e.g.*, 50-50-50), and random combinations (*e.g.*, 70-20-50). A SoftMax activation function is set as the final hidden layer in all networks to facilitate multi-class classification.

4. RESULTS AND DISCUSSION

We assess the outcomes of fault classification using two criteria: training duration and classification test accuracy (C_{Acc}), as shown in (10). True negative (TN) and true positive (TP) show the correctly classified cases, while false positive (FP) and false negative (FN) show cases that were not.

$$C_{Acc}(\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (10)$$

After conducting the training and testing of the proposed models, we obtain the following results: Concerning the binary trees, the maximum deviance reduction method achieved the highest accuracy of 98.83%. We also remark that the increase of maximum splits can greatly influence performance until a certain threshold where it starts to decrease. Table 4 shows the best preforming max number of splits for each module in all the split creation methods. Training time appears to be relatively low, especially for configurations with fewer splits. For ensemble trees, they demonstrated higher accuracy, with AdaBoost achieving an accuracy of 99.18% using 300 maximum splits and 50 learners at a learning rate of 0.15, as seen in Table 4. This demonstrates the power of combining multiple learners. While training times were higher than for single trees, the accuracy gain was significant.

The k-NN algorithms performed reliably, with the city block distance metric achieving 98.52% accuracy, outdoing Euclidean distance at 98.25%. Training time was consistently short, as seen in Table 4. Neural networks achieved the highest raw accuracy, with a three-layer architecture using the Tanh activation function reaching 99.28%; however, these outcomes came at a substantial computational cost, with training times exceeding four hours, as seen in Table 4. The trade-off between accuracy and training time is evident. The results across all models indicate a strong capability of different machine learning techniques to achieve high accuracy in classification tasks; Table 4 highlighted the best performers, and Figures 3(a) and 3(b) is the confusion matrix for the top two overall.

Table 4. The best preforming for each method

Type	Hyperparameters	Accuracy % (Test)	Training time
Binary Tree	Split method: Gini's diversity index | Maximum splits: 250	98.57331572	12.65 s
	Split method: Twoing rule | Maximum splits: 300	98.38837517	24.78 s
	Split method: Maximum deviance reduction | Maximum splits: 250	**98.83751651**	**30.68 s**
Ensemble Trees	Bagging | Maximum splits: 250 | Number of learners: 40	98.81109643	4 min 19.21 s
	RUS-Boost | Maximum splits: 250 | Number of learners: 40	98.89035667	4 min
	Ada-Boost | Maximum splits: 300 | Number of learners: 50	**99.18097754**	**4 min 21.32 s**
k-NN	Distance metric: Euclidean | Number of neighbors: 7	98.25627477	32 s
	Distance metric: City block | Number of neighbors: 15	**98.52047556**	**3 min 1.83 s**
ANN	1 layer | 1st layer size: 50 | function: Sigmoid	99.07529723	32 min 57.55 s
	2 layers | 1st layer size: 50 | 2nd layer size: 100 | function: ReLU	98.96961691	56 min 7.32 s
	3 layers | 1st layer size: 50 | 2nd & 3rd layer sizes: 100 | function: Tanh	**99.28665786**	**4 h 12 min 1.63 s**

Figure 3. The confusion matrix for the best performers, (a) random forest and (b) ANN

Our results show that the classification methods employed in this work achieve high accuracy while using a significantly lower algorithm complexity compared to CNNs or high-density ANNs. Such characteristics would render them more appropriate for bearing fault diagnosis and maintenance on an industrial level. The data set, however, is limited in size and variety of defects and working conditions. For future tests, we plan to use other, more complex datasets and try to implement and test real-time bearing fault monitoring for the prognosis and diagnosis of electric vehicles.

5. CONCLUSION

In this paper, we trained and tested various machine learning methodologies, specifically ANNs, k-NNs, binary trees, and ensemble trees, for bearing fault classification using time series data extracted from vibration signals. The results show that each method has its pros and cons when it comes to accuracy, prediction speed, and training efficiency. Notably, the ANN approach, particularly with a three-layered architecture, achieved the highest accuracy in classifying bearing conditions, while the AdaBoosted Tree model demonstrated a commendable balance between accuracy and training time.

This comparative analysis points out the advantages of leveraging basic ML methods to enhance fault detection in bearing-reliant machinery, ultimately contributing to improved maintenance strategies and the longevity of machinery. Future research may focus on refining these models, testing other techniques such as SVM and kernel-based ML methods, and exploring their applicability in real-time monitoring systems. This will probably advance predictive maintenance practices across various industrial sectors. By integrating these advanced models into existing maintenance frameworks, industries can reduce downtime and optimize operational efficiency, leading to significant cost savings and increased productivity.

In future work, the integration of these methodologies in electric vehicles (real/simulation) may accelerate the development of better predictive maintenance systems, enhancing the reliability and efficiency of rolling components such as wheel bearings, motors, and gear systems. This interface may ultimately foster safer and more efficient electric vehicle performance on roadways, extending their operational lifespan and minimizing maintenance costs for users.

FUNDING INFORMATION

Authors state that this research received no funds from public or private organizations.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Raid Houssem Eddine Sellaoui	✓	✓	✓		✓		✓	✓	✓	✓				
Brahim Boulebtateche		✓						✓	✓	✓		✓		
Salah Bensaoula		✓								✓		✓		

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, RHES, upon reasonable request.

REFERENCES

[1] A. Bankova, "Investigation of the qualitative dependence between the character of wear and the mutual location of wearing supports," in *2022 International Conference on Communications, Information, Electronic and Energy Systems, CIEES 2022 - Proceedings*, 2022, pp. 1–4, doi: 10.1109/CIEES55704.2022.9990870.

[2] A. J. Bazurto, E. C. Quispe, and R. C. Mendoza, "Causes and failures classification of industrial electric motor," in *Proceedings of the 2016 IEEE ANDESCON, ANDESCON 2016*, 2017, pp. 1–4, doi: 10.1109/ANDESCON.2016.7836190.

[3] Y. Merizalde, L. Hernández-Callejo, and O. Duque-Perez, "State of the art and trends in the monitoring, detection and diagnosis of failures in electric induction motors," *Energies*, vol. 10, no. 7, p. 1056, 2017, doi: 10.3390/en10071056.

[4] L. Cui, Z. Jin, J. Huang, and H. Wang, "Fault severity classification and size estimation for ball bearings based on vibration mechanism," *IEEE Access*, vol. 7, pp. 56107–56116, 2019, doi: 10.1109/ACCESS.2019.2911323.

[5] G. M. Lim, D. M. Bae, and J. H. Kim, "Fault diagnosis of rotating machine by thermography method on support vector machine," *Journal of Mechanical Science and Technology*, vol. 28, no. 8, pp. 2947–2952, 2014, doi: 10.1007/s12206-014-0701-6.

[6] K. Bouaouiche, Y. Menasria, and D. Khalfa, "Diagnosis of rotating machine defects by vibration analysis," *Acta IMEKO*, vol. 12, no. 1, pp. 1–6, 2023, doi: 10.21014/ACTAIMEKO.V12I1.1438.

[7] S. Aziz, M. Ahmed, I. Abbas, S. Z. H. Naqvi, and M. U. Khan, "Vibration signal analysis towards early detection of machine faults," in *Proceedings - 2020 1st International Conference of Smart Systems and Emerging Technologies, SMART-TECH 2020*, 2020, pp. 89–93, doi: 10.1109/SMART-TECH49988.2020.00034.

[8] W. Kurniawan, K. M. Hangos, and L. Márton, "Fault isolation and estimation in networks of linear process systems," *Entropy*, vol. 25, no. 6, p. 862, 2023, doi: 10.3390/e25060862.

[9] C. S. Jing *et al.*, "Fault detection and isolation for complex system," in *AIP Conference Proceedings*, 2017, vol. 1865, p. 70010, doi: 10.1063/1.4993392.

[10] A. Abbaspour, S. Mokhtari, A. Sargolzaei, and K. K. Yen, "A survey on active fault-tolerant control systems," *Electronics (Switzerland)*, vol. 9, no. 9, pp. 1–23, 2020, doi: 10.3390/electronics9091513.

[11] Y. A. Ghobashy, M. A. Geliel, and M. El Sengaby, "Industrial application of fault detection and fault isolation using artificial neural networks," in *ACCS/PEIT 2017 - 2017 Intl Conf on Advanced Control Circuits Systems and 2017 Intl Conf on New Paradigms in Electronics and Information Technology*, 2017, vol. 2018-February, pp. 48–59, doi: 10.1109/ACCS-PEIT.2017.8302999.

[12] M. Cerrada, G. Zurita, D. Cabrera, R. V. Sánchez, M. Artés, and C. Li, "Fault diagnosis in spur gears based on genetic algorithm and random forest," *Mechanical Systems and Signal Processing*, vol. 70–71, pp. 87–103, 2016, doi: 10.1016/j.ymssp.2015.08.030.

[13] T. W. Rauber, F. De Assis Boldt, and F. M. Varejão, "Heterogeneous feature models and feature selection applied to bearing fault diagnosis," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 1, pp. 637–646, 2015, doi: 10.1109/TIE.2014.2327589.

[14] Z. Su, B. Tang, Z. Liu, and Y. Qin, "Multi-fault diagnosis for rotating machinery based on orthogonal supervised linear local tangent space alignment and least square support vector machine," *Neurocomputing*, vol. 157, pp. 208–222, 2015, doi: 10.1016/j.neucom.2015.01.016.

[15] M. Alonso-Gonzalez, V. G. Diaz, B. Lopez Perez, B. Cristina Pelayo G-Bustelo, and J. P. Anzola, "Bearing fault diagnosis with envelope analysis and machine learning approaches using CWRU dataset," *IEEE Access*, vol. 11, pp. 57796–57805, 2023, doi:

10.1109/ACCESS.2023.3283466.
[16] Y. Chen, Q. Chen, and R. Wang, "Bearing fault diagnosis based on vibration envelope spectral characteristics," *Applied Sciences (Switzerland)*, vol. 15, no. 4, p. 2240, 2025, doi: 10.3390/app15042240.
[17] H. Qiu, J. Lee, J. Lin, and G. Yu, "Wavelet filter-based weak signature detection method and its application on rolling element bearing prognostics," *Journal of Sound and Vibration*, vol. 289, no. 4–5, pp. 1066–1090, 2006, doi: 10.1016/j.jsv.2005.03.007.
[18] K. Lee, J. K. Kim, J. Kim, K. Hur, and H. Kim, "CNN and GRU combination scheme for bearing anomaly detection in rotating machinery health monitoring," in *1st IEEE International Conference on Knowledge Innovation and Invention, ICKII 2018*, 2018, pp. 102–105, doi: 10.1109/ICKII.2018.8569155.
[19] Y. Zhang, J. Wu, H. Xie, R. Hua, and Q. Li, "NASA space station rolling bearings anomaly detection based on PARA-LSTM model," *International Journal of Sensor Networks*, vol. 44, no. 1, pp. 49–61, 2024, doi: 10.1504/IJSNET.2024.136334.
[20] L. Magadán, F. J. Suárez, J. C. Granda, F. J. DelaCalle, and D. F. García, "A robust health prognostics technique for failure diagnosis and the remaining useful lifetime predictions of bearings in electric motors," *Applied Sciences (Switzerland)*, vol. 13, no. 4, p. 2220, 2023, doi: 10.3390/app13042220.
[21] S. Y. Li, H. A. Li, L. M. Tam, and C. S. Chen, "A smart system for an assessment of the remaining useful life of ball bearings by applying Chaos-based health indicators and a self-selective regression model," *Sensors*, vol. 23, no. 3, p. 1267, 2023, doi: 10.3390/s23031267.
[22] M. Rahmaty, "Machine learning with big data to solve real-world problems," *Journal of Data Analytics*, vol. 2, no. 1, pp. 9–16, 2023, doi: 10.59615/jda.2.1.9.
[23] L. Sindayigaya and A. Dey, "Machine learning algorithms: a review," *International Journal of Science and Research (IJSR)*, vol. 11, no. 8, pp. 1127–1133, 2022, doi: 10.21275/sr22815163219.
[24] E. Karimov, "Binary tree," in *Data Structures and Algorithms in Swift*, Berkeley, CA: Apress, 2020, pp. 77–86.
[25] P. Cunningham and S. J. Delany, "K-nearest neighbour classifiers - a tutorial," *ACM Computing Surveys*, vol. 54, no. 6, pp. 1–25, 2022, doi: 10.1145/3459665.
[26] N. Ali, D. Neagu, and P. Trundle, "Evaluation of k-nearest neighbour classifier performance for heterogeneous data sets," *SN Applied Sciences*, vol. 1, no. 12, 2019, doi: 10.1007/s42452-019-1356-9.
[27] B. A. Paya, I. I. Esat, and M. N. M. Badi, "Artificial neural network based fault diagnostics of rotating machinery using wavelet transforms as a preprocessor," *Mechanical Systems and Signal Processing*, vol. 11, no. 5, 1997, doi: 10.1006/mssp.1997.0090.
[28] D. Gangadia, "Activation Functions: experimentation and comparison," in *2021 6th International Conference for Convergence in Technology, I2CT 2021*, Apr. 2021, pp. 1–6, doi: 10.1109/I2CT51068.2021.9417890.
[29] S. Sharma, S. Sharma, and A. Athaiya, "Activation functions in neural networks," *International Journal of Engineering Applied Sciences and Technology*, vol. 04, no. 12, pp. 310–316, 2020, doi: 10.33564/ijeast.2020.v04i12.054.
[30] D. Sacerdoti, M. Strozzi, and C. Secchi, "A comparison of signal analysis techniques for the diagnostics of the IMS rolling element bearing dataset," *Applied Sciences (Switzerland)*, vol. 13, no. 10, p. 5977, 2023, doi: 10.3390/app13105977.

BIOGRAPHIES OF AUTHORS

Raid Houssem Eddine Sellaoui received his M.Sc. degree in automation and industrial computer science from May 8th 1945 University in 2021. He is currently pursuing a Ph.D. degree in automation control at Badji Mokhtar Annaba University Algeria. His research interests include fault diagnosis, monitoring, machine learning, and their integration in electrical/hybrid vehicles. He can be contacted at email: raid-houssem-eddine.sellaoui@univ-annaba.dz.

Brahim Boulebtateche received his five-year State Engineer Degree (Diplôme d' Ingénieur ´d'Etat) with honors from National Polytechnic School of Algiers, Algeria (1984), his MPhil from Bradford University, Bradford, England, in 1988, and his PhD from University of Badji Mokhtar Annaba (UBMA), Annaba, Algeria, in 2007, all in electrical and electronic engineering. Since 1988, he has been with Badji Mokhtar University, Annaba (UBMA), where he is currently teaching in the Department of Electronics and he is a member of Laboratoire d'Automatique et Signaux de Annaba, (LASA). His current research interests include computer vision, machine learning, robotics and intelligent control systems. He can be contacted at email: brahim.boulebtateche@univ-annaba.dz.

Salah Bensaoula received his Engineer Degree of State from National Polytechnique of Algiers in 1983. The DEA degree from Clermont-Ferrand, France, and the Doctorate Degree from Saint-Etienne University, France, in 1984 and 1987, respectively. His main interests of research include fault detection and isolation in industrial systems and man-machine communication. He can be contacted at email: salah.bensaoula@univ-annaba.dz.