Ball bearing monitoring using decision-tree and adaptive neuro-fuzzy inference system

Riadh Euldji [1] , Mouloud Boumahdi[2], Mourad Bachene[2] and Rafik Euldji [2]

[1] *Laboratory of Automation and Industrial Diagnostic, University of Djelfa,* Algeria
[2] *Dept. mechanical engineering, University of Medea, Algeria*
r.euldji@univ-djelfa.dz

Abstract. This study aims to provide a methodology that relies on the combination of the following approaches: the decision tree, the neural network, and the fuzzy logic to monitor the evolution of bearing degradation. Data collected from the vibratory signals generated from the tests carried out on ball bearings mounted in an experimental fatigue platform, are used. The decision tree method is applied to select the most relevant monitoring indicator, which will be used to develop an Adaptive Neuro-Fuzzy Inference System (ANFIS). The training and test data required for model development have been classified according to the following states: normal, abnormal, and dangerous. These were defined from two thresholds: alert threshold and danger threshold. Then, the ANFIS model is trained from the indicators selected by the decision tree to predict the behaviour of the bearing in operation. The results confirm the effectiveness of the proposed approach for monitoring the health of ball bearing.

Keywords: Condition monitoring, Decision tree, ANFIS.

1 INTRODUCTION

For several decades and until today, vibration analysis continues to be the primary tool for analysing the behaviours of rotating machines. This approach aims to assess the state of health of a machine in real time and to transform a set of raw data collected on the monitored machine, using a data mining approach, into health indicators whose extrapolation over time makes it possible to offer support for the decision-making.

In the monitoring of a rotating machine, several problems can be encountered, such as: (1) choice and configuration of degradation state indicators; (2) estimation of the remaining operating time before the total degradation of the rotating element; (4) predict their future behaviour; (5) extraction of decision rules; (6) maintenance decision-making; (7) unavailability of experts in the expertise of rotating machines.

This work will contribute to overcoming the difficulties encountered when monitoring the condition of a rotating machine. It focuses on monitoring the behaviours of a ball bearings.
In the state of the art, several techniques have been proposed in order to predict the future behavior of ball bearings, e.g., Artificial neural network networks [1-4], support vector machines [5-6], decision tree [7-9], ANFIS [10].

this paper aims to propose a methodology to model the prediction of the behaviour of ball bearings in operation. This methodology relies on the application of the decision tree, and ANFIS on a set of real data.

The rest of this article is organized as follows. In Section 2, we present a methodology based on a decision tree and ANFIS approaches. Section 3 focuses on the data collection. In Section 4, we present the results obtained from applying the proposed approaches to a dataset. Conclusions are enclose in Section 5.

2 METHODDOLOGY

2.1 Decision tree

A decision tree is a classification method. It aims to extract information contained in data by using classification algorithms. The construction of this tree requires the definition of the features and the classes which form the dataset. The classification algorithm allows to choose the most important feature by using the criteria entropy and Gain ratio. These criteria are defined as follows:

Entropy used to select input variable

$$Info(T) = -\sum_{i=1}^{k} \left|C_j\right| (|T|)^{-1} \, log_2(C_j(T)^{-1}) \quad (1)$$

$$Gain\,(X_i,T) = Info\,(T) - Info(X_i,T) \quad (2)$$

Gain ratio to select splitting attributes

$$GR(X_i,T) = Gain\,(X_i,T)\,(Split\ info\,(X_i,T))^{-1} \quad (3)$$

Where $X=\{X_1,X_2,...,X_i, \dots,X_n\}$ is the features set, n is the number of features, $C=\{C_1,C_2,..., C_j, \dots,C_k\}$ is the classes set, k is the number of classes, $|C_j|$, $j=1, 2,..., k$ is the number of examples belonging to the class, T is the set of training examples, and $|T|$ is the total number of examples.

The feature chosen is the one that has the great gain ratio compared to the other features.

In order to build the desired decision tree, an algorithm must be used to classify the instances. Among the algorithms available, there are ID3 [11] and C4.5 [12]. The latter is the most widely used decision tree induction algorithm developed by Ross Quinlan. In this study, the J48 classification algorithm, a more developed version of C4.5, implemented in the WEKA software is used.

2.2 Adaptive neuro- fuzzy inference system (ANFIS)

ANFIS is a hybrid learning algorithm that incorporates fuzzy logic and artificial neural networks (ANN) in order to give enhanced prediction. The fuzzy system represents the reasoning part, while the neural networks represents the learning part. ANFIS uses the fuzzy if-then rules involving premise and consequent parts of Sugeno-type fuzzy inference.

To describe this system, we assume that the examined fuzzy inference system has two inputs x and y and one output *f*. For a first-order Sugeno fuzzy model, a common rule set with two fuzzy if–then rules is defined as:

1. If x is A1 and y is B1, then f1 = p1x + q1y + r1(4)

2. If x is A2 and y is B2, then f2 = p2x + q2y + r2 (5)

Where *p1, p2, q1, q2, r1* and *r2* are linear parameters in the consequent part and *A1, A2, B1* and *B2* are nonlinear parameters. The corresponding equivalent ANFIS architecture for two-input first-order Sugeno fuzzy model with two rules is shown in Fig. 1. The architecture of the ANFIS system consists of five layers, namely the fuzzy layer, product layer, normalized layer, de-fuzzy layer and total output layer.

Different layers of ANFIS have different nodes. Each node in a layer is either fixed or adaptive.

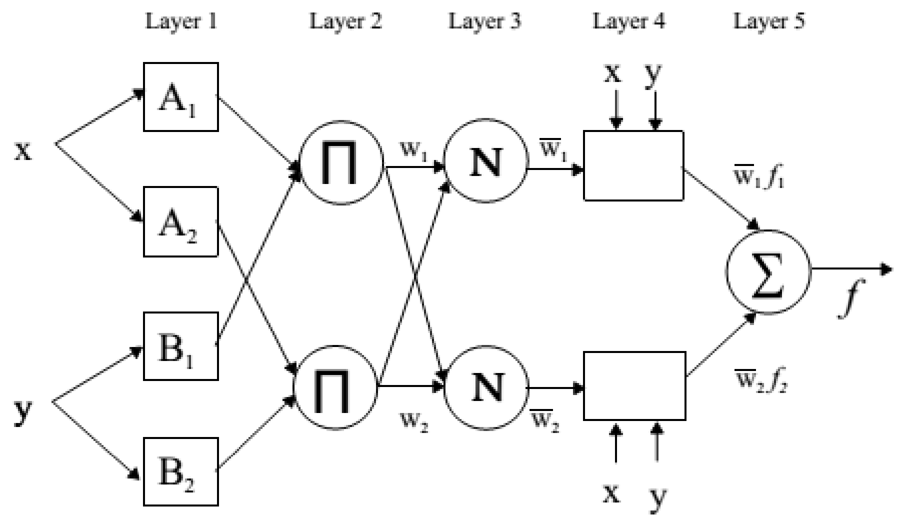

Fig. 1. ANFIS general structure.

3 DATA COLLECTION

To apply the proposed methodology, a data set was collected from the vibratory signals generated from the tests carried out on ball bearings mounted in an experimental fatigue platform, namely PRONOSTIA. This platform dedicated to test and validate bearings fault detection, diagnostic and prognostic approaches. The main objective of this platform is to provide real experimental data that characterize the degradation of ball bearings along their operational life [13].

The Nature of the data is vibration data measured by the vibration sensors during the time. Its features are namely: root mean square, kurtosis, Skewness, Peak, k factor, and the crest factor. Accounting, 2803 experimental data point samples were used to train the model. The mathematical description of each variable is presented in table 01

4 RESULTS AND DISCUSSION

4.1 Feature selection

There are three types of vibration analysis: time-domain analysis, frequency-domain analysis, and time–frequency analysis. The time domain analysis is a statistical analysis directly related to the time signal itself. In this analysis, several features are used in the vibratory follow-up of the bearing. We can quote to that end: root mean square, kurtosis, Skewness, Peak, k factor, and the crest factor (see Table 1). These features assess the state of global functioning of the machine but did not locate the defect. The presence of a defect can be detected if a feature exceeds a predetermined threshold.

In order to select the relevant feature, a decision tree algorithm, namely J48 is applied to the data set. This set contains information of three classes: normal, abnormal, and dangerous with six features (See Table 1). The classes were defined from two thresholds: alert threshold, and danger threshold, as follows: The alert threshold is the multiplication of the first measurement by two, which indicates the start of degradation, and the danger threshold is a multiplication of the first measurement by ten, which indicates the level of danger.

Table 1. Features used in the vibratory follow-up of the bearing.

Features	Description
Peak	$x_p = \max\lvert x(n)\rvert$
Root mean square	$x_{rms} = \sqrt{\frac{\sum_{n=1}^{N}\left(x(n)\right)^2}{N}}$
Kurtosis	$x_K = \frac{\sum_{n=1}^{N}(x(n)-p_1)^4}{(N-1).STD^4}$
Skewness	$x_{sks} = \frac{\sum_{n=1}^{N}(x(n)-p_1)^3}{(N-1).STD^3}$
Crest Factor (CF)	$x_e = \frac{x_p}{x_{rms}}$
K Factor(KF)	$KF = x_p * x_{rms}$

Figure 2 shows a decision tree constructed from a training data set. According to this figure, it can be seen that the root mean square, as a root node, appears to be more reliable than the other features. This indicates that the root mean square is the most relevant feature for making a decision in the anomaly detection process. After getting the most relevant feature, we will use them to build the ANFIS model.

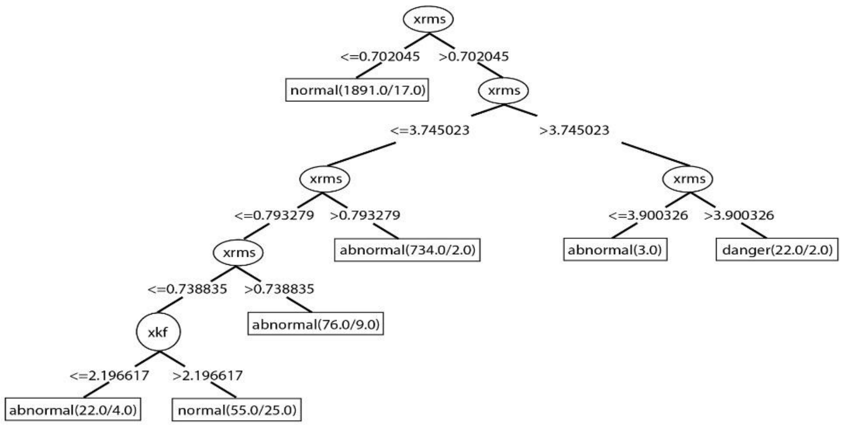

Fig. 2. Decision tree.

Table 2 shows the ranges of classification rate 0.97–0.98 and kappa statistics 0.95–0.89 which a classification rate of 1 means a perfect modelling and a kappa statistic of 0.7 or higher indicates a good statistics correlation.

Table 2. Performance of decision tree.

	Training data	Testing data
Classification rate	97.8951 %	98.1132 %
Kappa statistic	0.9514	0.8924
Mean absolute error	0.0216	0.0339
Number of instances	2803	1802

4.2 Prediction using ANFIS

From the previous results, the ANFIS model is trained using only the root mean square as a feature to predict the behavior of the ball bearings in operation.

The ANFIS training was performed with the time series of xrms. A xrms (t + 6) prediction is performed by using xrms(t), xrms(t − 6), xrms (t − 12), xrms (t − 18) data as input, which corresponds to past values.
Fig.3 and Fig.4 show that the feature root mean square presents a significant trend with regard to the evolution of the vibration level. The RMSE described in Table 3 is used to measure the prediction performance of the ANFIS.

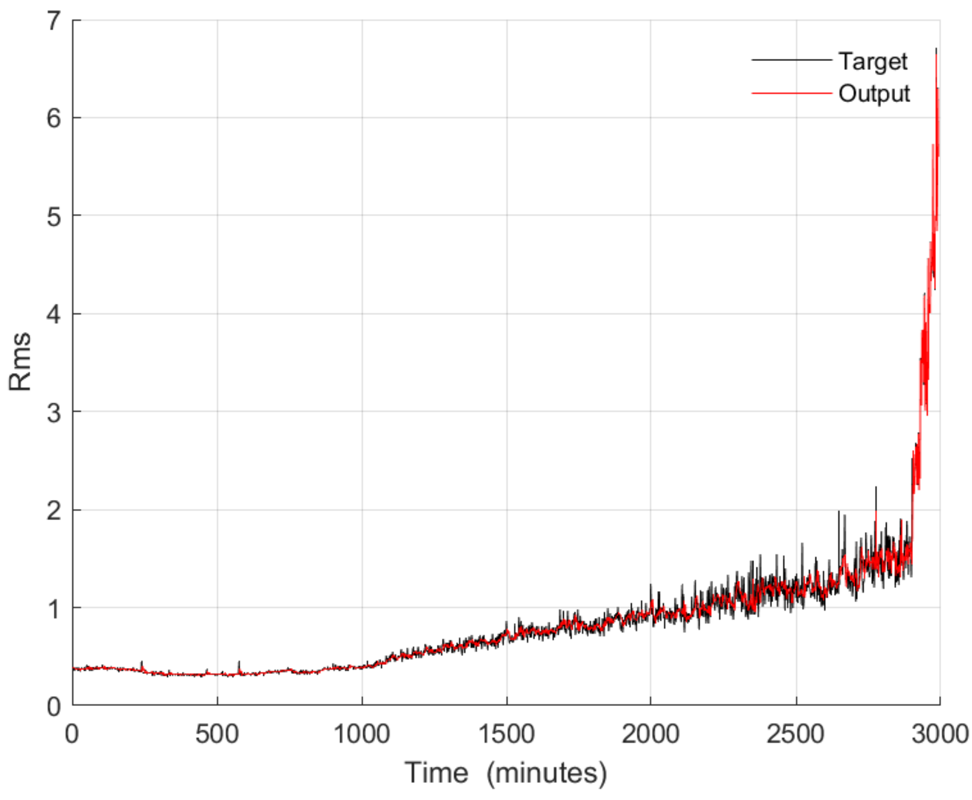

Fig. 3. Training data, experimental and predicted values of root mean square.

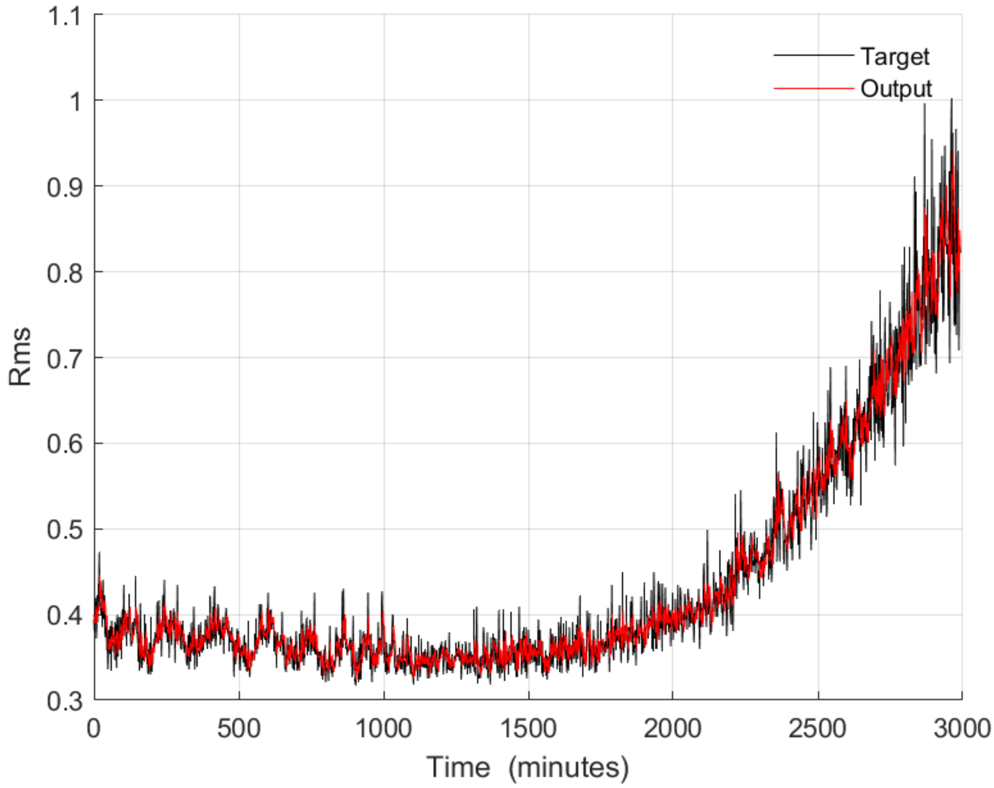

Fig. 4. Testing data, experimental and predicted values of root mean square.

Table 3. Performance of Anfis.

	Description	Training data	Testing data
MSE	$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2$	0.0072	0.0010
RMSE	$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}$	0.0850	0.0425

Figure 5 shows a good correlation among the ANFIS predictions and the experimental data (R^2 = 0.99149 for the training data and R^2 = 0.96752 for the testing data).

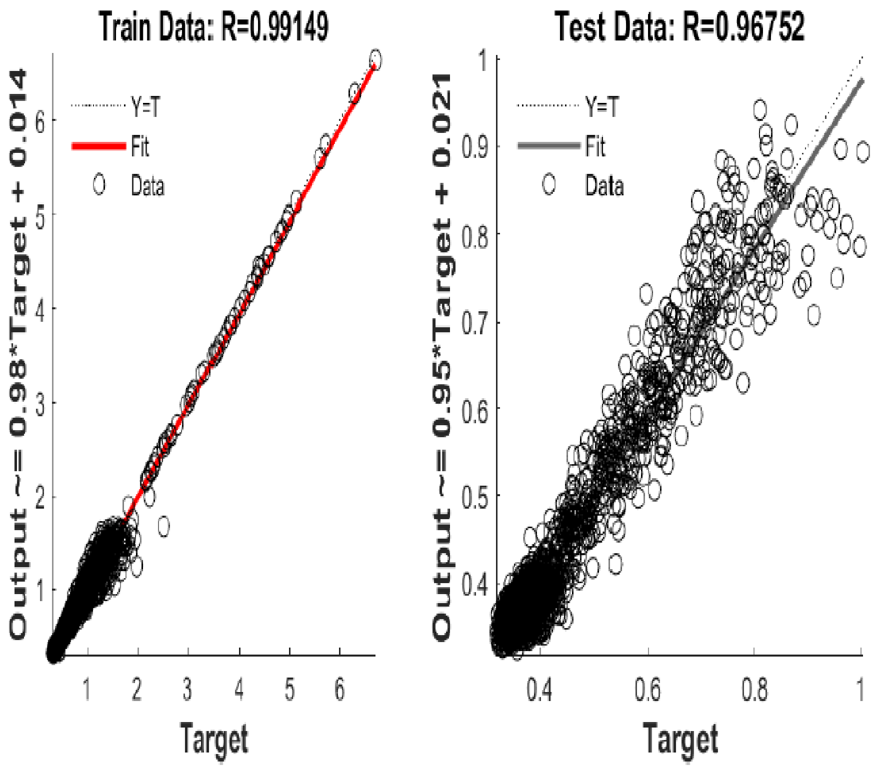

Fig. 5. Experimental and predicted values of root mean square.

5 CONCLUSION

In this study, we proposed a methodology to predict the behavior of ball bearings in operation. It relies on the vibration analysis and the application of the decision tree, and ANFIS on a set of real data collected from PRONOSTIA platform.

Two data set were used. The first dataset includes 2803 samples used for training, and the second includes 1804 samples used for testing. The data for each set were

classified into three state: normal, abnormal, and danger. The application of the decision tree algorithm on theses set allowed to classify the states of ball bearing perfectly, the ranges of classification rate 0.97–0.98 and kappa statistics 0.95–0.89. These results indicate that the feature RMS is the relevant feature to detect de degradation of the ball bearing. Then, the ANFIS model is trained using only the root mean square as a feature to predict the behavior of the ball bearings in operation. The value of R^2 indicate a good correlation among the ANFIS predictions and the experimental data.

References

1. Samanta, B., Al-Balushi, K.R.: Artificial neural network based fault diagnostics of rolling element bearings using time-domain. Mechanical Systems and Signal Processing, 17, 317-328 (2003).
2. Ali, J.B., Fnaiech, N., Saidi, L., Chebel-Morello B., Fnaiech, F.: Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals, Appl. Acoust., 89, 16-27 (2015),.
3. Unal, M., Onat, M., Demetgul, M., Kucuk, H. : Fault diagnosis of rolling bearings using a genetic algorithm optimized neural network, Measurement, 58, 187-196 (2014).
4. Chen, C., Vachtsevanos, G.: Bearing condition prediction considering uncertainty: an interval type-2 fuzzy neural network approach, Rob. Comput. Integr. Manuf., 28 (4), 509-516 (2012).
5. Rojas, A., Nandi, A.K.: Practical scheme for fast detection and classification of rolling-element bearing faults using support vector machines, Mech. Syst. Sig. Process., 20 (7), 1523-1536 (2006).
6. Sugumaran, V., Muralidharan, V., and Ramachandran, K.I.: Feature selection using decision tree and classification through proximal support vector machine for fault diagnostics of rolling bearing, Mechanical systems and signal processing, 21, 930–942 (2007).
7. Yang, B.S., Lim, D.S., and Tan, A.C.C.: VIBEX: an expert system for vibration fault diagnosis of rotating machinery using decision tree and decision table, Expert Systems with Applications, 28, 735–742 (2005).
8. Karabadji, N., Seridi, H., Khelif, I, Azizi, N., and Boulkroune, R.: Improved decision tree construction based on attribute selection and data sampling for fault diagnosis in rotating machines, Engineering Applications of Artificial Intelligence, 35, 71-83 (2014).
9. Sugumaran, V., and Ramachandran, K.I.: Automatic rule learning using decision tree for fuzzy classifier in fault diagnosis of roller bearing, Mechanical Systems and Signal Processing, 21, 2237–2247 (2007).
10. Ertunc, H.M., Ocak, H., Aliustaoglu, C.: ANN-and ANFIS-based multi-staged decision algorithm for the detection and diagnosis of bearing faults, Neural Comput. Appl., 22 (1) 435-446 (2013).
11. Quinlan, J. R.: Induction of decision trees. Machine Learning, 1, 81-106 (1986).
12. Quinlan, C4.5, J.R.: Programs for machine learning, Morgan Kaufmann Publishers, (1993).
13. Nectoux, P., Gouriveau, R., Medjaher, K., Ramasso, E., Chebel-Morello, B., Zerhouni, N., & Varnier, C.: PRONOSTIA: An experimental platform for bearings accelerated degradation tests. In IEEE International Conference on Prognostics and Health Management, PHM'12. IEEE Catalog Number: CPF12PHM-CDR, 1-8 (2012).