Application of artificial neural networks to diagnostics of fluid-film bearing lubrication

E P Kornaeva\textsuperscript{1}, A V Kornaev\textsuperscript{2,3}, Yu N Kazakov\textsuperscript{2} and R N Polyakov\textsuperscript{2}

\textsuperscript{1}Department of Information Systems, Orel State University named after I.S. Turgenev, 95 Komsomolskaya Street, Orel, 302026, Russian Federation
\textsuperscript{2}Department of Mechatronics, Mechanics and Robotics, Orel State University named after I.S. Turgenev, 95 Komsomolskaya Street, Orel, 302026, Russian Federation
\textsuperscript{3}E-mail: rusakor@inbox.ru

Abstract. The paper deals with the problem of monitoring the complicated hydromechanical system of a rotary machine with fluid-film bearings. The prospect of accurate recognition of changes in the lubrication system, which are expressed in the appearance of air bubbles in the lubricant, is investigated. The monitoring of the state is carried out by means of high-speed measurements of shaft vibrations, lubricant pressure supply and other parameters. Measurement data are transmitted for analysis to an artificial neural network to recognize the state of the system. The developed neural network has demonstrated high recognition accuracy of more than 98 \%. Some recommendations on the measurement results processing and the neural network settings are represented in the paper.

1. Introduction

Application of artificial neural networks (ANNs) for pattern recognition and diagnostics is a modern trend in engineering. Machine learning methods have become an alternative to classical deterministic methods in the field of diagnostics of rotary machines.

A method of pattern recognition and classification based on convolutional neural networks for rotary machinery is presented in [1]. The proposed method is based on signal processing in the form of images. Examples of several types of artificial neural networks for the state classification and defects detection of a planetary gear bearing assembly are given in work [2]. The relevance of using deep learning networks to classify data of the vibration amplitudes is established. In [3], the application of artificial neural networks with block diagnostics of a bearing assembly system is considered. The dataset passes through several blocks which leads to a greater generalization of data on the state of the system. It allows processing a large amount of information. Authors demonstrate that the method is effective, but requires improvements for applications in practice. In the article [4], the authors demonstrated the effectiveness of artificial neural networks for image processing. A method of the vibration signal converting into an image and application of artificial neural networks to classification of the state and fault diagnostics. In [5], the authors considered a method for recognizing bearing failures using a neuro-fuzzy system. It was established that such a model of the system allows fairly accurately classify various types of faults. The article [6] presents a method for processing data on the state of rotor systems using a convolutional neural network and data on thermal radiation. In the course of the work, an artificial neural network was...
trained using data from a thermal imager provided in the form of images. It is established that the use of this method of monitoring rotor systems is relevant and has high accuracy.

The works above are primarily related to the study of rotary machines with rolling bearings. In this paper, we consider the prospects of using artificial neural networks to recognize the state of the fluid-film bearing lubrication system. The classification of the state is carried out according to a complexly detectable feature of air bubbles presence in the lubricant.

2. Mathematical model
The classification problem under study relates to a supervised learning [7]. To develop a classifier program, a labeled dataset is needed for training. The classification principle can be considered as an example of two classes. In the case of many classes, classification can be implemented using so called “one vs. all” principle. Assumed that discrete quantity taking two values \( Y = \{0;1\} \) is a response variable. Continuous variables with values \( u_{ik} \) are considered as factors \( U_k \). The task is to approximate this data. In the simple case, the logistic regression can be used [7]:

\[
H_{\theta}(u) = (1 + \exp(-z(u)))^{-1},
\]

where \( z = z(u) \) is a function that determines the boundary between two classes when \( z(u) = 0 \).

The function \( H_{\theta}(u) \) is continuous with values in the interval \([0;1]\). Then relation of the \( i^{th} \) object to specific class according to hypothesis (1) can be determined using the following rule: «if \( H_{\theta}(u_i) \geq 0.5 \), then \( y_i = 1 \), else \( y_i = 0 \) ».

The boundary between classes can be described with polynomial of some degree. The specific polynomial function determines at the stage of testing of the model using a fragment of dataset. The procedure is called cross validation [7, 8].

For convenience of calculations, the boundary between classes is transformed into a polynomial of the first degree by replacing the terms of the original polynomial with new factors. For example, in case of two factors, the boundary in the form of a polynomial of higher degree “d” is transformed into a polynomial of the first degree using following algorithm (the “GNU Octave” programming language syntax is used [8]):

```matlab
for i = 1:d
    for j = 0:i
        U(:, end+1) = (u1.^((i-j)).*u2.^j);
    end
end
```

The values of each factor are recorded in columns in the matrix \( U = (u_{ik}) \), including the first unit column with the values of bias \( U_0 \). For the problem under consideration, the approximation function of classes is presented in the form of a complicated function, which is determined using an artificial neural network. The ANN architecture is represented in figure 1. The input layer includes \( N_{inp} + 1 \) neurons with \( N_{inp} \) input factors. The only hidden layer includes \( N_{hid} + 1 \) neurons. A hidden layer neuron summation result is a linear combination of input factors with bias. A sigmoid activation function (1) is used. The number of output neurons corresponds to the number of classes \( N_{out} \). The “softmax” activation function [7] is used in output neurons:

\[
H_i^{out} = \frac{\exp(z_i)}{\sum_{i=1}^{N_{out}} \exp(z_i)}.
\]
The goal of “softmax” function (2) is that the sum of all output values is always equal to one: \( \sum_{i=1}^{N_{out}} H_i^{out} = 1 \). Then the output values \( H_i^{out} \) can be interpreted as the probability of belonging to the specific class \( i \).

Figure 1. ANN architecture for solving the classification problem.

Unknown weights for each layer \( l \) are presented as matrices \( \theta_l^{(l)} = \left( \theta_{pq}^{(l)} \right)_{l=1,2} \). Index \( p \) corresponds to the neuron number in the current layer \( l \) and index \( q \) corresponds to the neuron number in the previous layer \( l-1 \). The hidden layer neurons activation function (1) should be calculated as follows:

\[
H_{hid} = H_{\theta} \left( U \left( \theta_{hid} \right)^T \right),
\]

where \( H_{hid} \) is a matrix with \( [Q \times N_{hid}] \) elements, \( U \) is a matrix with \( [Q \times (N_{inp}+1)] \) elements and \( \theta_{hid} \) is a weights matrix with \( [N_{hid} \times (N_{inp}+1)] \) elements.

The “softmax” activation function (2) on the output layer is calculated as a linear combination of the hidden layer neurons calculation results, and the bias:

\[
H_{out} = H_{\theta} \left( \bar{H}_{hid} \theta_{out} \right),
\]

where \( H_{out} \) is a matrix with \( [Q \times N_{out}] \) elements, \( \bar{H}_{hid} \) is a matrix with \( Q \times (N_{hid}+1) \) elements and \( \theta_{out} \) is a weights matrix with \( [N_{out} \times (N_{hid}+1)] \) elements.

The weights calculation procedure is an unconstrained optimization problem for the loss function. Taking into account the regularization procedure [7] the loss function for the classification problem can be represented as (the Einstein summation notation is used):

\[
J(\theta_{hid}, \theta_{out}) = - \frac{1}{Q} \sum_{i=1}^{Q} \left( \bar{y}_{ip} \ln H_{out}^{pi} + (1 - \bar{y}_{ip}) \ln (1 - H_{out}^{pi}) \right) + \frac{\lambda}{2Q} \left( \theta_{pq}^{hid} \theta_{pq}^{hid} + \theta_{pq}^{out} \theta_{pq}^{out} \right) \rightarrow \text{min},
\]

where \( H_{out}^{pi} \) are the components of the matrix \( H_{out} \), \( \bar{y}_{ip} = \left( \bar{y}_{ip} \right) \), \( y_{ip} = (y_{ip} = p) \) are converted values of known values of belonging to classes from the set \( \{0;1\} \), \( i = 1,Q \), \( y_{ip} \) are known values of belonging to classes from the set \( \{1,...,N_{out}\} \), \( p = 1,N_{out} \) are the class numbers and \( \lambda \) is a regularization coefficient.
The further procedure for determining unknown weighting factors for a large number of factors is carried out numerically by the gradient descent method. The initial values of weights $\theta^{(1)}$ are determined using uniform distributed random value $R \sim (0;1)$. The components of the gradient of the loss function (6) can be represented as:

$$
\frac{\partial J}{\partial \theta_{pq}^{out}} = \frac{1}{Q} \left( \delta_{pq}^{out} \tilde{H}_{iq}^{hid} \right),
$$

$$
\frac{\partial J}{\partial \theta_{p0}^{out}} = \frac{1}{Q} \left( \delta_{p0}^{out} \tilde{H}_{iq}^{hid} + \lambda \theta_{p0}^{out} \right),
$$

(6)

where $\delta_{ip}^{out} = H_{ip}^{out} - \tilde{Y}_{ip}$, $\tilde{Y}_{ip} = (\tilde{y}_{ip})$, $\tilde{y}_{ip} = (y_i \equiv p)$, $p = 1, Nout$, $q = 1, Nhid + 1$, $i = 1, Q$.

$$
\frac{\partial J}{\partial \theta_{p0}^{hid}} = \frac{1}{Q} \left( \delta_{p0}^{hid} u_{iq} \right),
$$

$$
\frac{\partial J}{\partial \theta_{pq}^{hid}} = \frac{1}{Q} \left( \delta_{pq}^{hid} u_{iq} + \lambda \theta_{pq}^{hid} \right),
$$

(7)

where $\delta_{iq}^{hid} = \delta_{ip}^{out} \delta_{pq}^{out} , i = 1, Q, p = 1, Nhid + 1$, $\delta_{ip}^{hid} = H_{ip}^{hid} (1 - H_{ip}^{hid})$, $p = 2, Nhid + 1$, $i = 1, Q$, $q = 1, Ninp + 1$.

The iterative procedure is performed using the standard function “fminunc ()” from the “GNU Octave” library. The arguments of the “fminunc ()” function are the loss function gradient components (7) - (8) and the value of the loss function (6) calculated in the previous step.

3. The test rig

The test rig for fluid-film bearings study includes a bronze bearing of 20 mm wide and 40.2 mm diameter (see figure 1). The average bearing clearance is 100 μm. The steel shaft’s length is 380 mm and it’s diameter is 40 mm. The hollow shaft mass is 0.6 kg. The bearing is lubricated with water, which is supplied under pressure through the cover of the bearing housing.

![Figure 2](image_url)
The scheme of measuring and control system is represented on Figure 3. The following modules by ‘National Instruments’ are used as analog-to-digital converters and digital-to-analog converters: ‘NI 9205’, ‘NI 9215’, ‘NI 9481’ and ‘NI 9269’. The modules are integrated in the ‘NI CDAQ-9178’ chassis.

The ‘NI 9205’ module is used to convert analog signals with the pressure sensor ‘KPT5-3’, the voltage divider for measuring the contact resistance of the fluid film [9], the flow meters ‘YF-S201 ’ for cold and hot fluid circuits and the motion sensors’ AP2100A- C-05.05.1 ’. The input signal range is ± 10 V.

The ‘NI 9215’ module is used to convert analog signals from the ‘Altivar 312’ frequency controller to obtain the energy consumption data of the ‘ELTE TMPE3 12/2’ electro motor. The Relay output module ‘NI 9481’ is used to control the ‘Smart SM88634’ distribution servo valves, which automatically enable one of the four cases of lubricant supply (see figure 2 (a) and figure 3).

The ‘NI 9269’ module is a digital-to-analog converter. In this test rig, it is used to control the fluid flow rate in two hydraulic circuits with ‘Burkert 2835’ and ‘EV260B’ servo valves, as well as to control the electric motor by means of the ‘Altivar 312’ frequency inverter. The module generates a signal in the range of ± 10 V.

4. Results and discussion

The training and testing of an artificial neural network to recognize one of four cases of lubricant supply with or without air bubbles (see figure 2, b) was carried out based on the results of a physical experiment. Seven parallel tests were performed for each lubricant supply case. Thus, the total number of tests was 28. The duration of each test was 90 seconds. The shaft was gradually accelerated to 4200 rpm, then gradually decelerated and stopped. The servo valves were configured so that the lubricants pressure supply and flow rates were approximately the same for all methods of lubricant supply. The sampling rate of the sensors was 1 kHz. Such a sampling rate is quite sufficient for representing trajectories of oscillations of the shaft at the highest rotational speed.

After the experiment, the measurement data were processed using the Savitzky-Golay filter to remove high-frequency noise and smooth the signal. A fragment of the smoothed results of measuring the shaft vibrations in the form of trajectories is presented in figure 4.
Figure 4. The shaft’s vibration trajectories (μm) in a time range 60-70 s for four cases of lubrication: pure fluid (line 1), mixture of pure fluid and aerated fluid (lines 2 and 4) and aerated fluid (line 3).

Figure 5 demonstrates the frequency response for the vertical oscillations of the shaft. It can be seen from the results (see figures 4, 5) that the presence or absence of air bubbles in the lubricant does not uniquely affect the shape of the trajectories and the amplitude of the vibrations.

Further, to identify patterns and diagnose the condition of the bearing lubrication system, training and testing of an artificial neural network were performed. The mathematical model of the network is presented in Section 2 of the present paper, and its architecture is represented in figure 1. The dataset was divided in a proportion of 75:15:15 for training, validation and testing, respectively.

The artificial neural network training was carried out both for data processed using a filter, and for raw data, and for their combination. Testing results demonstrated clearly that signal filtering unambiguously increases the accuracy of recognizing.

Figure 5. The shaft’s vibration frequency response (Hz) for the vertical vibrations (μm) for four cases of lubrication: pure fluid (line 1), mixture of pure fluid and aerated fluid (lines 2 and 4) and aerated fluid (line 3).
The additional testing with various settings of the network and with various datasets was carried out as follows. The effect of the number of measured quantities, frame length, number of samples and number of hidden neurons was investigated (see figure 6). When varying one of the values, the rest were fixed at an average level. The average levels were as follows. The average frame length was 400. This means that 400 signals from each sensor were used to generate one sample. The average number of training samples was 1200. And the average number of hidden neurons was 40.

![Figure 6. The ANN’s accuracy testing.](image)

The results represented in figure 7 demonstrate that the number of measured quantities (using 1, 2 or 3 sensors) has the greatest influence on the accuracy of recognition. When measuring one quantity (rotor vibrations magnitude in one of the directions), the recognition accuracy was 70-80%. When measuring two quantities (oscillations in two mutually perpendicular directions), the accuracy increases to 85-90%. When measuring three values (two components of the oscillations and pressure), the accuracy reaches 95-98%.

5. Conclusions
The application of feed forward neural networks with one hidden layer to monitor the state of the fluid-film bearing lubrication system has shown successful results. As an input data, the results of measuring the rotor vibrations and the lubricant supply pressure were used. As a response, the presence or absence of air bubbles in the lubricant was calculated. Testing results demonstrated a high diagnostic accuracy of up to 98%. The greatest impact on the accuracy of the network had the number of measurement sensors. It was also found that the recognition accuracy is significantly higher if the measurement data is filtered from noise.

Acknowledgments
The work has been carried out within the project No. 05.607.21.0303 "Development of an intelligent monitoring technology and a prototype of the hardware and software complex for safety of energy complex facilities" supported by the Federal Target Program "Research and development in priority areas for the development of the scientific and technological complex of Russia for 2014 - 2020" by the Ministry of Science and Higher Education of Russian Federation (unique identifier of the research RFMEFI60719X0303) and with the project No. 16-19-00186 supported by the Russian Scientific Research Foundation (the test rig development in section 3 of the present paper). The authors gratefully acknowledge this support. The authors also express graduate to Alexey Rodichev and Alexander Fetisov for the help with the test rig modification.

References
[1] Zhang J, Sun Y, Guo L, Gao H, Hong X and Song H 2019 A new bearing fault diagnosis method based on modified convolutional neural networks Chinese J. of Aeronautics 1350 1–9
[2] Zhao D, Wang T and Chu F 2019 Deep convolutional neural network based planet bearing fault classification Computers in Industry 107 59–66
[3] Zhu Z, Peng G, Chen Y and Gao H 2019 A convolutional neural network based on a capsule network with strong generalization for bearing fault diagnosis 323 62 – 75
[4] Hoang D-T and Kang H-J 2019 Rolling element bearing fault diagnosis using convolutional neural network and vibration image Cognitive Systems Research 53 42–50
[5] Abdelkrim C, Meridjet M S, Boutasset N and Boulanouar L 2019 Detection and classification of bearing faults in industrial geared motors using temporal features and adaptive neuro-fuzzy inference system Helikon 5 1–11
[6] Li Y, Du X, Wan F, Wang X and Xiao Q-B 2019 Rotating machinery fault diagnosis based on convolutional neural network and infrared thermal imaging Chinese J. of Aeronautics 1384 1–12
[7] Goodfellow I, Bengio Yo and Courville A 2016 Deep Learning (MA: MIT Press) p 802
[8] Eaton J W, Bateman D, Hauberg S and Wehbring R 2015 GNU Octave 4.0 Reference Manual 1/2 (Hong Kong: Samurai Media Limited) p 536