Mecánica Computacional Vol. XXI, pp. 1912–1921
S. R. Idelsohn, V. E. Sonzogni and A. Cardona (Eds.)
Santa Fe–Paraná, Argentina, October 2002

COMPARISON OF OPTIMIZATION TECHNIQUES OF NEURAL NETWORKS TRAINING FOR FAULTS DIAGNOSTIC OF ROTATING MACHINERY

Darley F. A. Santiago[*] and Robson Pederiva[+]

[*] UNICAMP / Faculty of Mechanical Engineering
CEFET-PI / Centro Federal de Educação Tecnológica do Piauí
P. O. Box: 6051, Campinas, São Paulo, Brazil, ZIP: 13083-970
e-mail: darley@fem.unicamp.br

[+] UNICAMP / Faculty of Mechanical Engineering
P. O. Box: 6051, Campinas, São Paulo, Brazil, ZIP: 13083-970
e-mail: robson@fem.unicamp.br

Key words: Neural Networks, Faults Diagnostic, Rotating Machinery

Abstract. *Recently sophisticated vibration monitoring techniques have been available to be used in the monitoring and diagnostics of complexes rotating machinery. Among them, can relate the artificial intelligence techniques as neural networks, fuzzy logic, expert systems and so on. The neural networks are tools that have woken up a lot of interest on researchers in the recent years. They let the monitoring on-line of predictive maintenance aiming the minimization of the time between the receiving of the information and the diagnosis of the problem. This paper shows the ability and feasibility of the application of different optimization techniques of neural networks training in the diagnostic of faults inserted in the rotating machinery. In the experimental setup are inserted the following faults: defect electric, mechanical looseness, unbalance + mechanical looseness and unbalance. Several architectures of neural networks implemented with the Matlab software were trained with different optimization techniques to provide the best architecture to diagnostic of four faults inserted in the experimental setup. Results show that the neural networks can be effectively used in the diagnostic of faults inserted in the experimental set-up with a high performance and that the Levenberg-Marquardt optimization technique is faster than gradient descent and gradient descent with momentum for practical problems.*

1 INTRODUCTION

The maintenance predictive is a science that use several kinds of data to determine the condition of the machine and to predict a fault before it occurs. The aims of maintenance predictive are in general, decrease the downtime, lower maintenance costs and improved security. For today's sophisticated systems of machinery, predictive maintenance has become the most reliable method for monitoring and diagnosis of faults. There are many different types of methods that may be used, including oil analysis, vibration analysis and temperature and pressure monitoring. For many years, vibration analysis has been widely accepted as the most reliable method for predicting machinery problems. The vibration signals are used for rotating machinery condition monitoring, fault diagnosis and severity estimation. Fault detection and diagnosis is generally accepted to occur in three stages[1]:

- Detection – has a fault occurred?
- Identification – where is the fault?
- Diagnosis – why has the fault occurred?

The importance of detection and diagnosis of faults in machinery has expanded considerably in the past decade due to the increased complexity of plant equipment and high costs associated with failure and shutdown. Fault recognition normally requires a detailed analysis of machinery signals to identify specific fault patterns. Traditionally, this is performed through visual inspection by experienced personnel using spectrum analysis or associated signal processing methods. However, these methods are usually costly and inefficient in the some cases. As an alternative to conventional fault diagnostic methods, artificial intelligence techniques are being introduced to assist in fault diagnosis[2,3]. Among them, can relate neural networks, fuzzy logic, expert system and so on. The neural networks are tools that have woken up a lot of interest on researchers in the last years. They let the monitoring on-line of predictive maintenance aiming the minimization of the time between the receiving of the information and the diagnostic of the problem[4]. There are many applications of neural networks in the diagnostic of mechanical faults[5-10].

This work shows the ability and feasibility of the application of different optimization techniques of neural networks training in the diagnostic of faults inserted in the rotating machinery. The acquisition of data were generated on a laboratory rotor-rig.

In the experimental setup were inserted the following faults: defect electric, mechanical looseness, unbalance + mechanical looseness and unbalance. Several architectures of neural networks implemented with the Matlab software were trained with different optimization techniques to provide the best architecture to diagnostic of faults inserted in the experimental setup. Results show the application of neural networks for faults diagnostic of rotating machinery using real data, as well as its theoretical and practical aspects of implementation.

2 NEURAL NETWORKS

Basically, the artificial neural network consists of neurons, simple processing elements, which are activated as soon as their inputs exceed certain thresholds. The neurons are arranged in layers which are connected so that the signals at the input are propagated through the network to the output[11]. A neuron is a processing unit of information indispensable for the operation of the neural network. The Figure 1 shows the model of the neuron.

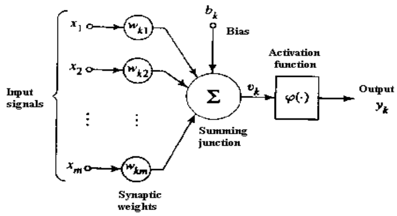

Figure 1: Model of the neuron

The neuron can be represented mathematically as follows:

$$v_k = \sum_{j=1}^{m} w_{kj} x_j \tag{1}$$

$$y_k = \varphi(v_k + b_k) \tag{2}$$

where,
x_j : are the input signals;
w_{kj} : are the synaptic weights;
v_k : is the linear combiner output;
b_k : is the bias;
$\varphi(.)$: is the activation function;
y_k : is the output signal of the neuron.

The bias externally applied in the neuron, have the effect of increase or decrease a liquid input of the activation function, when it is positive or negative, respectively. The activation functions can be three types: limiar function, piecewise-linear function and sigmoid function.

A feedforward neural network contains one or more layers[12]. Three types of layers may exist in a neural network: an input layer, an output layer an one or more hidden layers if necessary. These networks are named of multilayer perceptron (MLP). In this work we chose the MLP because it provides a complex nonlinear mapping between the input and the output and is found to be simple to implement. Several techniques for training neural networks are available in the literature, the most common of them is backpropagation. The backpropagation algorithm consists of two steps through of layers of the network: a forward step, the propagation and a backward step, error back-propagation. In the forward step, one pattern of input or signal is propagated through the layers of the network as long as the synaptic weights are kept constants. This result of the network output or response due to input pattern is subtracted of the desired response and the error is so propagated backward through the network. During this step the synaptic weights are updated. In other words, the algorithm changes individually the synaptic weights until the goal error pre-determined to be reached by network. The goal error is defined through the quadratic mean sum squared error $(e.q.m)$ as follows:

$$e.q.m. = \frac{1}{N}\sum_{k=1}^{N}(y_k - a_k)^2 \tag{3}$$

where, a_k is the desired response.

There are many optimization techniques for neural networks training using the backpropagation algorithm. Among them, can relate the gradient descent method, gradient descent with momentum, conjugate gradient method (Fletcher-Reeves, Polak-Ribiere), quasi-Newton method (Broyden-Fletcher-Goldfarb-Shanno - BFGS), Levenberg-Marquardt method and so on. It is very difficult to know which training algorithm will be the fastest for a given problem. It will depend on many factors, including the complexity of the problem, the number of data points in the training set, the number of weights and biases in the network, and the error goal. The two first methods are often slow for practical problems. In general, on networks which contain up to a few hundred weights the Levenberg-Marquardt algorithm will have the fastest convergence. This advantage is especially noticeable if very accurate training is required. The quasi-Newton methods are often the next fastest algorithms on networks of moderate size. For more details about the optimization techniques for neural networks training, see[12,13].

3 RESULTS AND DISCUSSION

This section presents some results obtained through the implementability and training of several neural networks architectures with the backpropagation algorithm, using real data as network input features. The Matlab Neural Networks Toolbox 3.0 was used for neural networks implementability. The real data (patterns) used for training, test test and I II were generated in the experimental setup, shown in Figure 2. The experimental setup consists of an electric motor 0.5 HP, 110 V AC, one rotor fixed in the motor shaft and supported at both ends for two identical rolling element bearings. The faults were inserted in the experimental setup separetely, and were collected the vibration signals (velocity) using an accelerometer mounted in the right rolling bearing in vertical direction.

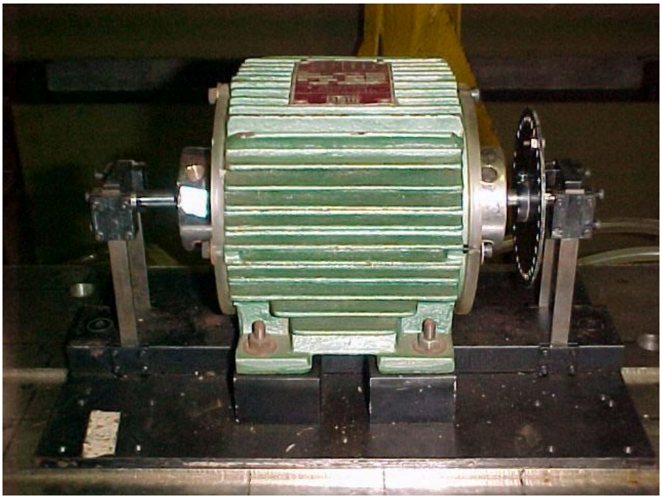

Figure 2: Experimental Setup

As follows is described some individual features of each fault type, as well as show a spectrum of the measured signal for each fault inserted in the experimental setup.
Defect electric – This type of fault is further common in rotating machinery and can be caused basically by eccentricity of the rotor and unbalance voltage line. In the second case, this defect can appear in the spectrum with frequency equal a twice the frequency of the supply line and the dominant plane is the radial with low amplitude. The spectrum in this signal collected during the stage of data acquisition is shown in Figure 3.
Mechanical looseness – This type of fault was inserted in the experimental setup by loosening one of the four bolts between the electric motor and the basis. The spectrum in these signal collected during the stage of data acquisition is shown in Figure 4.
Vertical unbalance + Mechanical looseness – The unbalance was inserted in the experimental setup by addicting a mass of about 7.8 g at any end of the rotor. The mechanical looseness appears combined the unbalance. The spectrum in this signal collected during the stage of data acquisition is shown in Figure 5.
Horizontal unbalance – This fault was inserted in the experimental setup by addicting a mass of about 2.5 g at any end of the rotor. The spectrum in this signal collected during the stage of data acquisition is shown in Figure 6.

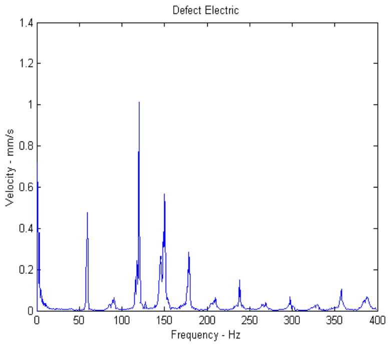

Figure 3: Spectrum of Defect Electric

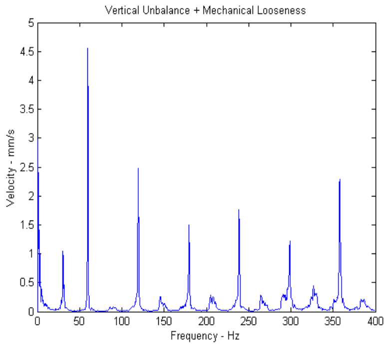

Figure 5: Spectrum of Vertical Unbalance + Mechanical Looseness

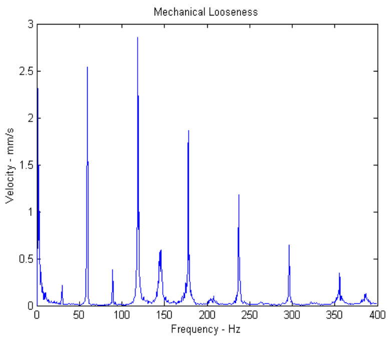

Figure 4: Spectrum of Mechanical Looseness

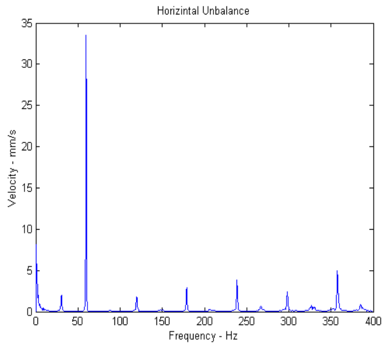

Figure 6: Spectrum of Horizontal Unbalance

During the acquisition of the vibration signals (patterns) regarding the first three faults, the accelerometer was mounted in the right rolling bearing in the vertical direction, and for the fourth fault the accelerometer was mounted in the right rolling bearing in the horizontal direction. Eighty such vibration signals were collected, being twenty signals for each of the faults inserted. The rotor speed was about 3600 rpm. The sample frequency was of 400 Hz and a line frequency of 60 Hz.

The real data set used for training, test I and test II of the architectures of neural networks implemented to diagnostic the four faults inserted in the experimental setup was split of the following way: Forty spectrums (patterns) with data set of training of the network (ten patterns for each type of fault); twenty patterns with data set of test I (five patterns for each type of fault) and twenty patterns with data set of test II (five patterns for each type of fault).

As input features of network were chosen the signal amplitudes of each fault (FRF – Frequency Response Function) in the frequencies of 1xRPM, 2xRPM, 3xRPM, 4xRPM, 5xRPM and 6xRPM, where RPM is a rotor speed. As output features of the network were used the following target values (activation level): 1000 (defect electric), 0100 (mechanical looseness), 0010 (vertical unbalance + mechanical looseness) and 0001 (horizontal unbalance). A value equal to 1 in a dimension is a symbol of pressure of a particular fault, while a value equal to 0, means its absence. The aim of the application of neural networks in the real case is investigating its efficiency and feasibility with tool of faults diagnostic of rotating machinery. During the stages of implementability, training, test I and test II of the backpropagation neural network were considered the following training features:

- Input and hidden layers : hyperbolic tangent function activation;
- Output layer : linear function activation;
- Goal error : 1e-5.

Several architectures of neural networks were trained with different numbers of neurons in the input and hidden layers.

Table 1 shows the results obtained for neural networks training with different optimization techniques using an architecture 6x5x4. In all cases it presented a 100% of rate success for set data test I and test II.

Table 1 – Comparison of Optimization Techniques of Neural Networks Training – Real Case.

Optimization Technique	Epochs Number	Training Time (s)	Success Rate (%) Test-I Data	Success Rate (%) Test-II Data
Gradient Descent	99673	1577	100	100
Gradient Momentum	64528	867	100	100
Fletcher-Reeves	418	13	100	100
Polak-Ribiere	294	10	100	100
quasi-Newton	182	9	100	100
Levenberg-Marquardt	15	3	100	100

The Table 1 gives some convergence times examples for the several algorithms (optimization techniques of training) on one particular problem. Ten different test run were made for each training algorithm in a Pentium III 800 MHz PC to obtain the average numbers shown in the table. The training process using the gradient descent method was carried out with learning rate coefficient equal to 0.15 and to gradient descent with momentum method equal to 0.9. These values of learning rate coefficient and momentum were kept constant.

Other tests were carried out increasing the learning rate coefficient to 0.25 and the results showed decreasing of the training time and 100% of rate success for set data test I and test II.

During training is important to know that if the learning rate is made too large, the algorithm will become unstable and if the learning rate is set too small, the algorithm will take a long time to converge. Generally, learning rate is a small number (0.05-0.9). The other way to decrease the convergence time avoiding the unstability, is by adopting a momentum term.

Momentum allows a network to respond not only to the local gradient, but also to recent trends in the error surface. Without momentum a network may get stuck in a shallow local minimum. With momentum a network can slide through such a minimum. Momentum term can be any number between 0 and 1. When the momentum is 0 the gradient descent with momentum method will become gradient descent method.

Table.1 shows that gradient descent method and gradient descent with momentum were slow for practical problem. On the other hand, the conjugate method, quasi-Newton method and Levenberg-Marquardt method were extremely fast in the faults diagnostic inserted in the experimental setup. The advantage of the Levenberg-Marquardt is that it converges faster around the minimum and gives more accurate results. Its only drawback is that it requires more memory than the backpropagation with momentum method. Figures 7-10 show satisfactory results obtained with the Levenberg-Marquardt method applied in the diagnostic of four faults inserted in the experimental setup using neural networks.

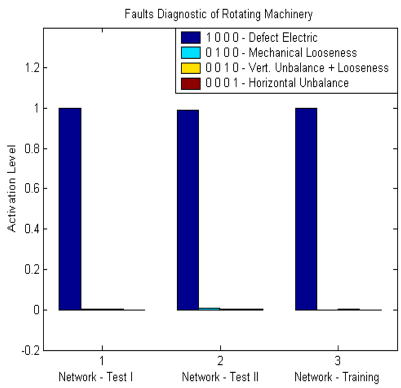

Figure 7: Comparison Between the Training, Test I and Test II Networks – Defect Electric

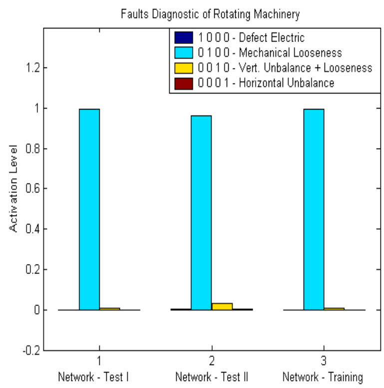

Figure 8: Comparison Between the Training, Test I and Test II Networks - Mechanical Looseness.

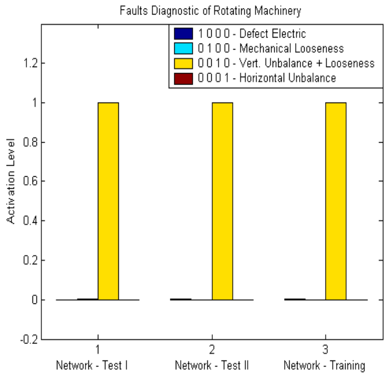

Figure 9: Comparison Between the Training, Test I and Test II Networks - Vertical Unbalance + Mechanical Looseness.

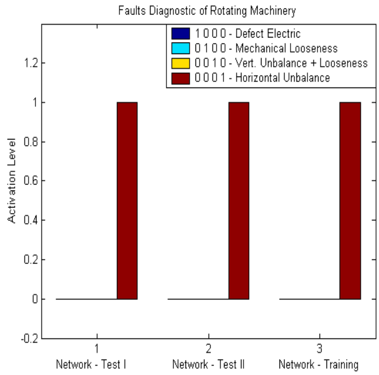

Figure 10: Comparison Between the Training, Test I and Test II Networks - Horizontal Unbalance.

4 CONCLUSIONS

Neural are tools that have woken up a lot interest on researchers in the lately years. They let the monitoring on-line of predictive maintenance aiming the minimization of time between receiving the information and the diagnostic of the problem.

Observing the results presented was seen that: from many theoretical and practical aspects that are related the one neural network design, the choice of a neural network architecture and of its training features not abide rules predefined; the knowledge and experience of the designer in regard to problem faced are more important. The definition stage is delicate, well involve, to be beyond the choice of neural network architecture, the obtaining of significant variables set to problem resolution. This obtaining involves, beyond variables identification that are related with the problem, the removing of variables not reliable to process, or whose use is not practicable for economical and technical reasons.

In addition, we observed that the sensitivity and time response of neural network in regard the other conventional techniques of faults diagnostic are important features and can be assessed and improved during the stages of implementability, training and tests of neural network.

Finally, we showed that using a real case, the ability and feasibility of the application of neural networks with tool further efficiency in the faults diagnostic of rotating machinery. Results show that neural networks can be effectively used in the diagnostic of faults inserted in the experimental setup with a high performance and that the Levenberg-Marquardt optimization technique is faster and gives more accurate results for practical problems.

5 REFERENCES

[1] Th. Amand, G. Heyen and B. Kalitventzeff, "Plant Monitoring and Fault Detection Synergy Between Data Reconciliation and Principal Component", *Computers and Chemical Engineering*, **25**, 501-507 (2001).
[2] C.-C. Lin and H.-P. Wang, "Performance Analysis of Rotating Machinery Using Enhanced Cerebellar Model Articulation Controller (E-CMAC) Neural Networks", *Computers Ind. Engineering*, **30(2)**, 227-242 (1996).
[3] D. C. Baillie and J. Mathew, "A Comparison of Autoregressive Modeling Techniques for Fault Diagnosis of Rolling Element Bearing", *Mechanical Systems and Signal Processing*, **10(1)**, 1-17 (1996).
[4] A. Lucifredi, C. Mazzieri. and M. Rossi, "Application of Multiregressive linear Models, Dynamic Kriging Models and Neural Network Models to Predictive Maintenance of Hidroeletric Power Systems", *Mechanical Systems and Signal Processing*, **14(3)**, 471-494 (2000).
[5] X. Wu., J. Ghaboussi and J. H. Garret-Jr, "Use of Neural Networks In Detection of Structural Damage", *Computers & Structures*, **42(4)**, 649-659 (1992).
[6] I. E. Alguidingue, A. L. Buczak and R. E. Uhrig, "Monitoring and Diagnosis of Rolling Element Bearings Using Artificial Neural Networks", *IEEE Transactions on Industrial Electronics*, **40(2)**, 209-217 (1993).
[7] M.-Y. Chow, R. N. Sharpe and J. C. Hung, "On The Application and Design of Artificial Neural Networks for Motor Fault Detection – Part II", *IEEE Transactions on Industrial Electronics*, **40(2)**, 189-196 (1993).
[8] T. I. Liu, J. H. Singonahalli and N. R. Iyer, "Detection of Roller Bearing Defects Using Expert System and Fuzzy Logic", Mechanical *Systems and Signal Processing*, **10(5)**, 595-614 (1996).
[9] C. Zang and M. Imregun, "Structural Damage Detection Using Artificial Neural Networks and Measured FRF Data Reduced Via Principal Component Projection", *Journal of Sound and Vibration*, **242(5)**, 813-827 (2001).
[10] N. S. Vyas and D. Satishumar, "Artificial Neural Network For Fault Identification In a Rotor-Bearing System", Mechanism and Machine Theory, **36**, 157-175 (2001).
[11] P. M. Frank and B. Kopen-Seliger, "Fuzzy Logic and Neural Network Applications to Fault Diagnosis", *Int. Journal of Aproximate Reasoning*, **16**, 67-88 (1997).
[12] S. Haykin, "Neural Networks: A Comprehensive Foundation", New Jersey, Prentice Hall (1999).
[13] H. Demuth and M. Beale, "Neural Network Toolbox User's Guide 3.0, MathWorks, Inc. (1998).