

## A multi-layer based procedure for detecting liver fibrosis

SMARANDA BELCIUG, MONICA LUPSOR

---

**ABSTRACT.** The aim of this paper was to assess the effectiveness of MLPs application to the liver fibrosis evaluation and to establish the comparison between the two MLP types, regarding their performances on this particular dataset.

*2000 Mathematics Subject Classification.* 92B20, 68T05.

*Key words and phrases.* neural networks, multilayer perceptron (MLP), backpropagation (BP), liver fibrosis, fibroscan.

---

### 1. Introduction

The progression of chronic liver diseases can be forecasted using as an indicator the hepatic fibrosis. The last stage of liver stiffness reaches cirrhosis and hepatocellular carcinoma. This is a very important issue because chronic diseases and cirrhosis are the causes of morbidity and mortality in many countries around the world. The data given by the World Health Organization in January 2004 shows that Egypt has 14,610 deaths, United States 14,003 deaths, Mexico 10,887 deaths, Romania 10,023 deaths, Japan 9,849 deaths.

Liver biopsy was for 60 years the gold standard diagnostic for assessing the progression of fibrosis [3]. This procedure is though an invasive and painful procedure, carrying a significant, although small risk of life-threatening complication [4]. A new technological discovery worldwide in the evaluation of hepatic fibrosis is the Fibroscan (Echosens, Paris, France). The fibroscan is a specially adapted ultrasound device that uses the principle of the one-dimension transient elastography (TE) for the assessment of liver stiffness.

Modern medical methods provide diagnosis accuracy for detecting hepatic fibrosis ranging from 22% to 71%[7]. One of them [6], concerning the "one against all" technique, used ROC (Receiver Operating Characteristic) analysis [5] to provide thresholds between different stages of liver fibrosis, by repeatedly grouping them into two-subgroups (multi-step procedure). This is a pretty slow technique since it is based on randomly splitting patients into two initial groups and then refining step-by-step the procedure until the right stage is correctly identified.

The aim of this paper is to use machine learning techniques such as the multilayer perceptron with 2 hidden layers (*3-layer perceptron*) and the multilayer perceptron with 3 hidden layers (*4-layer perceptron*) applied to the medical dataset, and statistically evaluate their performances.

---

*Received:* 19 June 2009.

## 2. Neural Networks

Neural Networks (NNs) are parallel computing devices consisting of many interconnected simple processors. Basically, NNs are trying to mimic the behavior of the human brain. Each of the processors from the network is quite simplistic, but a collection of them is a powerful computational tool.

NNs process information in a similar way that the human brain does. The network is composed of a large number of neurons (processing units) that are interconnected, and work in parallel to solve a specific problem. NNs learn by example and they cannot be programmed to perform a specific task. The examples must be selected carefully, otherwise useful time is wasted or even the network might be functioning incorrectly.

The most popular NN in use today is the *multi-layer perceptron* (MLP). Basically, a MLP represents a configuration of simple perceptrons, layered in a hierarchical structure forming a feed-forward topology. Technically, it consists of one or more hidden layers of perceptrons between the input and output layers. Let us notice that it is rare to use more than three hidden layers. Among the learning algorithms used in the training process the most common one is the *back-propagation (of-error)* algorithm (BP). MLP does generally work fairly well, but it tends to be slow when the training dataset is large [2], [9]. In this study we have used two MLP types: (a) *3-layer perceptron*, (b) *4-layer perceptron* and the BP algorithm as learning rule.

Mathematically, the network input of weighed signals received by a unit is given by the following formula:

$$net_j = w_0 + \sum_{i=1}^n w_i \cdot x_i,$$

where  $w_0$  is the biasing signal,  $w_i$  is the weight on the input connection  $ij$ ,  $x_i$  is the magnitude of signal on input connection  $ij$ , and  $n$  is the number of input connections to unit  $j$ .

The BP learning algorithm consists of the following steps:

1. *Forward pass*: The activity of the input units is determined by the network's external input. For all other units, the activity is propagated forward:

$$y_i = f_i\left(\sum_{j \in A_i} w_{ij} y_j\right).$$

2. *Error computation*: compute the difference between the desired output and actual output. Assuming that we are using the sum-squared loss, given by:

$$E = \frac{1}{2} \sum_o (t_o - y_o)^2,$$

the error for output unit  $o$  is simply  $\delta_o = t_o - y_o$ , where  $t_o$  represents the desired output and  $y_o$  the actual network output.

3. *Backward pass*: propagate error back through the network to recursively compute weight changes. We can expand the error of hidden unit in terms of its posterior nodes:

$$\delta_j = - \sum_{i \in P_i} \frac{\partial E}{\partial net_i} \frac{\partial net_i}{\partial y_j} \frac{\partial y_j}{\partial net_j}.$$

Of the three factors inside the sum, the first is just the error of node  $i$ . The second factor is:

$$\frac{\partial net_i}{\partial y_j} = \frac{\partial}{\partial y_j} \sum_{k \in A_i} w_{ik} \cdot y_k = w_{ij},$$

while the third factor is the derivative of node  $j$ 's activation function:

$$\frac{\partial y_j}{\partial net_j} = \frac{\partial f_j(net_j)}{\partial net_j} = f'(net_j).$$

Putting all the pieces together, we get the error of hidden unit:

$$\delta_j = f'_j(net_j) \sum_{i \in P_i} \delta_i w_{ij}.$$

4. *Learning updates*: the synaptic weights are updated using the results of the forward and backward passes.

### 3. The Dataset

743 consecutive patients with chronic HCV infection examined at the 3rd Medical Clinic, University of Medicine and Pharmacy "Iuliu Hatieganu" Cluj-Napoca, Romania, between May 2007 and August 2008 were prospectively included in this study. All of them had positive HCV-RNA in their serum and underwent percutaneous liver biopsy (LB) for grading and staging the diseases. All patients were referred to liver stiffness measurement (LSM) 1 day prior to LB. Besides the epidemiological, anthropometric and clinical parameters, the biological parameters were determined for all patients on the same day as LSM.

The study was approved by the local Ethical Committee of the University of Medicine and Pharmacy "Iuliu Hatieganu" Cluj-Napoca. The nature of the study was explained to the patients, each of whom provided written informed consent before the beginning of the study, in accordance with the principles of the Declaration of Helsinki (revision of Edinburgh, 2000).

Transient elastography (TE) was performed using the FibroScan device, which consists of a 5-MHz ultrasound transducer probe mounted on the axis of a vibrator. The vibrator generates a completely painless vibration (frequency 50 Hz and amplitude 2 mm) which is similar to a "flick", generating an elastic shear wave that propagates through the skin and the subcutaneous tissue to the liver. The velocity of the wave is directly related to tissue stiffness. As previously described [8], and as suggested by the provider of the instrumentation, we considered representative measurements 10 successful acquisitions with a success rate of at least 60% (the ratio of the number of successful measurements to the total number of acquisitions) and with an interquartile range (IQR) lower than 30% of the median value. IQR reflects the variability of the validated measures.

The database consists of 25 predictive attributes and the class diagnosis with 5 values (Metavir F score).

### 4. Results

#### Summary statistics

The main statistic parameters included in this analysis are the following:

- Mean;
- Median;
- Confidence interval;
- Standard deviation.

The *mean* is a particularly informative measure of the "central tendency" of the variable if it is reported along with its confidence intervals.

A measure of central tendency, the *median* of a sample is the value for which one-half (50%) of the observations, when ranked, will lie above that value and one-half will lie below that value.

The *standard deviation* (SD) is a commonly used measure of variation. The standard deviation of a population of values is computed as the square root of the corresponding variance.

The *confidence intervals* for the mean give us a range of values around the mean where we expect the "true" (population) mean is located (with a given level of certainty, e.g. 95%). More precisely, if we repeatedly calculated this interval from many independent random samples of the same size, then 95% of the intervals would, in the long run, correctly bracket the true value of the mean, or equivalently we would in the long run be correct 95% of the time in claiming that the true value of the mean is contained within the confidence interval. Table 1 below summarizes the above statistical parameters.

Table 1. NN classification performances over 100 computer runs

NN classification performances over 100 computer runs				
Mode	Mean	95 % confidence interval(mean)	Median	Std.Dev.
Training	0.7333	(0.72, 0.74)	0.7334	0.031
Testing	0.5431	(0.53, 0.55)	0.5496	0.024

The testing mean performance equaling 54% is consistent with the usual medical diagnosis results [4], [6], [8]. The main advantage in using the MLP technique is representing by: (a) a much faster diagnosing process; (b) different from the "one against all" technique, using the ROC analysis, which is 'step by step' procedure, it represents an overall diagnosing process.

**Normality test** Many types of statistical analysis are based on the assumption that the data are normally distributed. Thus, we have used both the *Kolmogorov-Smirnov* & *Lilliefors* test (which is applicable when the mean and the standard deviation are computed from the actual data), and the *Shapiro-Wilk*  $W$  test [1]. Table 2 below provides the corresponding results.

Table 2. Normality tests for NN performances over 100 computer runs

Normality tests for NN performances		
Mode	Kolmogorov-Smirnov $d$ / Lilliefors $p$ -level	Shapiro-Wilk $W$ / $p$ -level
Training	0.17 / > 0.20	0.951 / 0.17
Testing	0.16 / > 0.20	0.950 / 0.17

### Visualization

The box and whisker plot will summarize each variable by three components:

1. A central square to indicate central tendency or location (e.g. mean);
2. A box to indicate variability around this central tendency;
3. Whiskers around the box to indicate the range of the variable.

**Note.** If data follows the normal distribution, then  $\text{mean} \pm 1.96 \cdot \text{SD}$  represents the 95% confidence interval and, therefore, the corresponding data should fall within the whiskers.

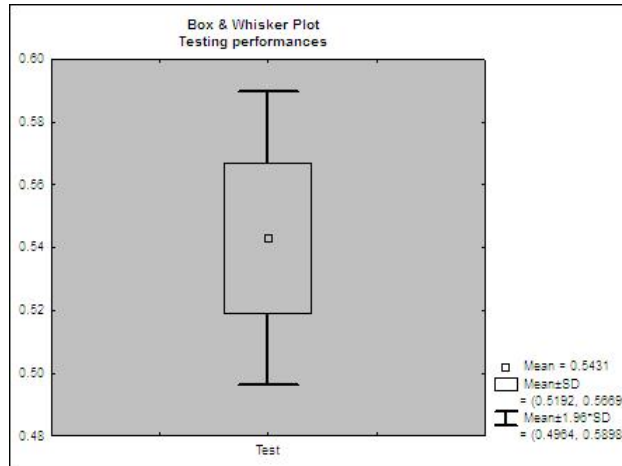


Fig. 1 Box and whiskers plot - testing performance

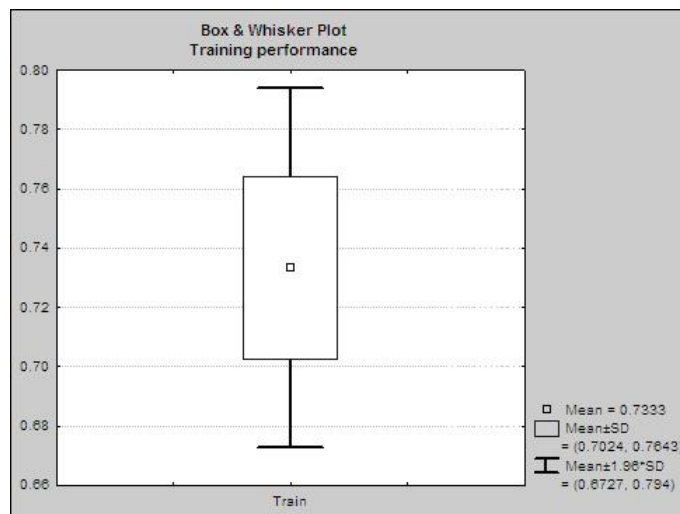


Fig. 2 Box and whiskers plot - training performance

### Scatterplot/Correlation

The scatterplot of the points related to both training and testing performances, together with the Pearson's  $r$  correlation coefficient, displayed in Fig. 3, show a negative correlation between the two performance parameters, that is, an increasing trend for training implies a decreasing trend for testing, which may indicate a slight over-learning trend.

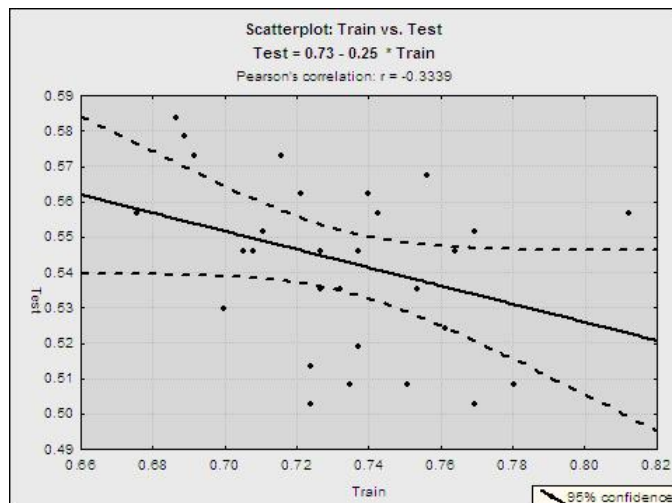


Fig. 3 Scatterplot of training/testing (negative correlation)

## 5. Conclusions

The aim of this paper was to assess the effectiveness of MLPs application to the liver fibrosis evaluation and to establish the comparison between the two MLP types, regarding their performances on this particular dataset. From this study we can conclude that the use of MLP in liver fibrosis evaluation is efficacious, comparable as accuracy with usual medical techniques, but much faster.

The statistical tests revealed that there is no significant difference between the two MLPs, that is we can consider only the 3-layer perceptron which is faster than the other option. This work can be deepened by considering the use of other machine learning techniques and find out the best classifiers.

## References

- [1] Altman, D. G., Practical statistics for medical research, Chapman and Hall,(1991).
- [2] Bishop, C., Neural Networks for Pattern Recognition, *Oxford University Press*, (2004).
- [3] Bravo, A.A., Sheth, S.G., Chopra, S., Liver biopsy. *N Engl J Med* , 344: 495-500, (2001).
- [4] Cadranet, J.F., Rufat, P., Degos F., Practices of liver biopsy in France: results of a prospective nationwide survey. For the Group of Epidemiology of the French Association for the Study of the Liver (AEEF). *Hepatology*; 32: 477-481, (2001).
- [5] Fawcett T. An introduction to ROC analysis. *Pattern Recognition Letters*; 27: 861- 874, (2006).
- [6] Lupsor, M., Badea, R., Stefanescu, H., Grigorescu, M., Sparchez, Z., Serban, A., Branda, H., Iancu, S., Maniu, A., Analysis of histopathological changes that influence liver stiffness in chronic hepatitis C. Results from a cohort of 324 patients. *J Gastrointestin Liver Dis.* 17(2): 155-163, (2008).
- [7] Nunez M., Serum Markers for Assessment of Liver Fibrosis in HIV-HCV Coinfected Patients, 57 Annual Meeting of the American Association of Liver Diseases, October 26-31, 2006, Boston, MA (2006).
- [8] Sandrin, L., Fourquet, B., Hasquenoph, J.M., et al., Transient elastography: a new noninvasive method for assessment of hepatic fibrosis. *Ultrasound Med Biol* 29: 1705-1713, (2003).
- [9] Zaknich A., *Neural Networks for Intelligent Signal Processing*, World Scientific, (2003).

(Smaranda Belciug) DEPARTMENT OF INFORMATICS, UNIVERSITY OF CRAIOVA,  
 AL.I. CUZA STREET, NO. 13, CRAIOVA RO-200585, ROMANIA, TEL. & FAX: 40-251412673

(Monica Lupsor) DEPARTMENT OF ULTRASONOGRAPHY 3RD MEDICAL CLINIC UNIVERSITY OF  
MEDICINE AND PHARMACY, CLUJ - NAPOCA

*E-mail address:* smaranda.belciug@inf.ucv.ro, mlupsor@yahoo.com