

Neural Approaches to the Diagnosis and Characterization of the Lyme Disease

G. Bianchi*, L. Buffirini*, P. Monteforte*, G. Rovetta*, S. Rovetta**, R. Zunino**

* Ist. Bruzzone Rheumatological Center, Dept. of Internal Medicine, Genoa University

** Department of Biophysics and Electronic Engineering, Genoa University

Abstract

Our primary goal was to develop and to study neural approaches to classification and rule extraction. We used for the necessary validation phases the Lyme database, recorded to study the Lyme borreliosis, by the Italian Working Group on Lyme disease (Ist. Bruzzone Rheumatological Center - Dept. of Internal Medicine of Genoa University). A procedure using a combination of statistical analysis and a standard supervised neural method obtained a very good classification (diagnosis) performance, but the resulting rules (characterization) were readable only in a rough form. A different approach was used to extract more readable rules. A binary-units network was trained by an information-theoretic learning rule (CCE). The result is a hierarchy of rules describing the classification by a sequence of yes/no questions, confirming the known diagnostic criteria to a satisfactory extent. The next step was to obtain a confirmation from an unsupervised method. An auto-associative network trained by back-propagation was used. The output of this method was a clustering of the patterns, and their interpretation led to the same conclusions of the CCE approach.

Keywords: Neural networks - Automatic diagnosis - Lyme disease - Rule extraction

1: Introduction

Medical diagnosis problems are often approached using pattern analysis and recognition techniques. Medical databases are generally available in a form that is suitable for easy translation into "patterns"; they also have been pre-processed by people who have a deep knowledge of the medical problem, thus eliminating the need for unsafe processing by non-expert researchers (that could imply information loss if not properly done). On the other hand, pattern-oriented approaches often need sound validation to be fully understood in all their strengths and weaknesses; this step of the study must be carried on with the aid of a real-world data set. So it is useful both to knowledge engineers and to medical researchers to have an interdisciplinary cross-check on their respective sides of the problem: pattern analysis techniques to solve a "real world" problem, and automated data study to extract information from a medical database.

Many kinds of techniques have been proposed to approach this problem. This study reports on a collection of works, each one dealing with a particular method or class of methods applied to classification and characterization of the same clinical problem: Lyme disease.

2: Lyme Disease and the Lyme Database

Automating the diagnosis and characterization processes for Lyme disease is a challenging task from both a knowledge-engineering and a medical viewpoint. Attempts to implement computer-aided processes are quite away from technical significance. This holds for most approaches from classical AI, which often adopt the expert system model (rule-based reasoning

programs) or feature some variation on the basic model. No one of them is completely satisfactory: average error rates range from 10% to 25% or even more, depending on how general the system aims to be (that is, on the number of diseases the system knows).

This situation is not very good, but in the case of Lyme disease additional difficulties arise specifically from the problem under consideration. Lyme borreliosis has been studied since the end of XIX century as a group of unrelated or weakly connected pathologies, but until 1977 it was not recognized as a single disease (first cases studied in 1975, coming from Old Lyme, Connecticut). This is because of the extreme variety of forms it can present.

LYME BORRELIOSIS
Caused by <i>Borrelia Burgdorferi</i> , brought by <i>Ixodes Dammini</i> , <i>Ixodes Ricinus</i> and other ticks
Found in specific geographic locations
Early phase (2 weeks since infection):
- ECM
- Non-specific symptoms (fever, weakness)
- Specific anticorpal response: IgM > 256:1 (standard threshold)
Middle phase (1 to 6 months):
- Cardiac symptoms
- Neurological symptoms
Late phase (1 year or more):
- Arthritis, responding to antibiotic therapy
- Specific anticorpal response: IgG > 256:1

Table 1 Description of Lyme disease

Its main symptom, the typical skin manifestation called Eritema Chronicum Migrans (ECM), is very peculiar, but quite rare. Other signs are very confusing, because they can easily lead to completely wrong diagnoses. In a first phase, for instance, the patient can complain only a persistent fever. In later phases, there can be other manifestations, some of which are typical (lymphadenosis benigna cutis, acrodermathitis chronica atrophicans) but some are aspecific or even misleading (arthritis, neurological and cardiological involvement). Only in a few cases the list of symptoms is clearly related to Lyme borreliosis, because generally only some of the most peculiar ones are present, and they may span over a time interval going from months to years, and sometimes they may even be neglected if Lyme disease is not suspected. Moreover, Lyme disease is caused by *Borrelia Burgdorferi*, a spirochaeta carried by different kinds of tick (*Ixodes Dammini*, *Ixodes Ricinus* and others), which bite is not painful, and therefore is not very apparent. This means also that it is present only in a few geographical areas, where there are animals parasited by ticks, ticks are parasited by *Borrelia Burgdorferi*, and humans have the opportunity to come in frequent contact with ticks (because statistically it is not so easy to be infected). As a consequence, Lyme disease is not only difficult to recognize, but also relatively rare to find.

The data used for our studies was the *Lyme* database, which contained data recorded in ten years (1983-1993) of clinical observations and laboratory tests performed by the Italian Working Group on Lyme disease (Ist. Bruzzone Rheumatological Center - Dept. of Internal Medicine of Genoa University). The database included 741 samples of patients, of which almost 200 were classified as Lyme-affected (*ill*); so the rest was a set of negative examples (*non-ill*).

3: Methods

Our main goal is to obtain a help in diagnosing and characterizing Lyme Borreliosis using automatic pattern recognition and classification methods, outperforming standard expert systems in learning capabilities and performance. This goal is twofold, in that we have two different objectives: diagnosis (classification of each patient as ill or non-ill) and characterization (description of criteria used by the classification task).

3.1: Feed-Forward Networks Trained by Backpropagation with Pre-Analysis and Compression

Diagnosis has been first approached by a supervised method, namely, a classical neural network trained by backpropagation [6]. The diagnostic performance thus obtained was not very satisfactory; a data analysis step was needed to pre-process the data set. This preliminary step has been carried on by training a standard network, and then analyzing the resulting classification rule with a statistical procedure, which gave as a result a selection of less important recordings and symptoms. Successively, another standard network was trained with an unsupervised method to compress these symptoms. This preprocessing partially "shrank" training patterns, reducing their (apparent) dimensionality but concentrating their information content. A final network trained on these "lighter" data obtained a very good classification performance, that is, it operated a correct diagnosis. However, the resulting rules were readable only in a rough form (useful/less useful variables), and thus characterization was poor [4]. The method is summarized in Fig. 1.

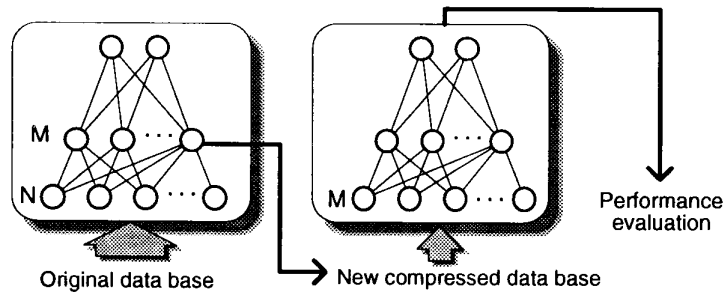


Fig. 1 Compression with a neural network

3.2: Feed-Forward Networks Trained by CCE Minimization

A good characterization can be achieved by extracting from a good classifier a set of meaningful rules. To obtain these readable rules, a totally different approach was used. A binary-units network was trained by minimization of conditional class entropy (CCE) [5], a training algorithm that not only tries to minimize classification error, but also to optimize the network layout by adding one hidden unit at a time, thus minimizing the network dimension.

Extracting rules from a binary network is a matter of attributing a true/false value to each unit; the significance of the rules obtained is ensured by the training algorithm, that also gives an importance ranking to each variable. Thus, the final result is a hierarchy of rules describing the

two classes (ill/non-ill) by a sequence of yes/no questions.

3.3: Auto Associative Feed-Forward Networks Trained by Back-Propagation

The next step was to obtain a confirmation from an *unsupervised* method. The chosen approach was an auto-associative five layers network trained by back-propagation (AABP) [3], which requires the same pattern as an input and as a target, and features an inner bottleneck, a central hidden layer with a small number of units, acting as a (non-linear) principal components analyzer. This unsupervised approach was elected because of its simple implementation with the same software tools used for the back-propagation. The output of this method is a clustering of the patterns, given by the values of the units in the bottleneck. The assumption for this cluster analysis is that the distribution of activations can provide hints of clusters in data. By applying the patterns to the input of the network, and reading the values on the central hidden layer, the bottleneck, we obtained one- and two-dimensional clusters (respectively from one- and two-unit hidden layers).

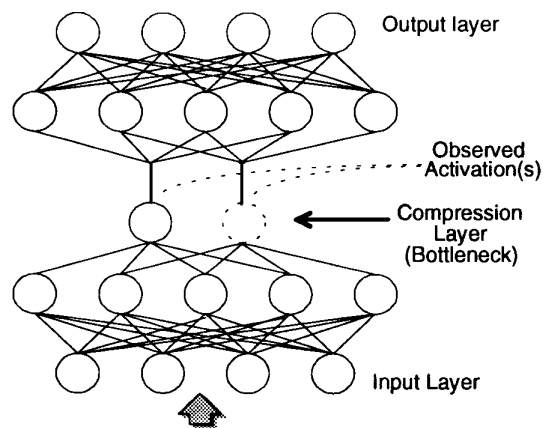


Fig. 2 Auto-associative feedforward network

4: Results and Comments

The diagnosis process, implemented with a neural network of the standard kind plus auxiliary compression, showed good error rates. The characterization was good from a performance standpoint, but characterization was not very good.

The ordered list of significant descriptors for Lb coincided with clinical expectation, including, for example, the presence of ECM or the progress of IgG among the most relevant indicators. The *overall error percentage* averages at about 8%, which appears quite satisfactory as compared to other automated or empirical approaches. These results show that the diagnostic performance is high if compared to other, more traditional methods such as "expert systems."

To judge the results of the characterization experiments, we need to observe the clusters obtained by each procedure. A good characterization is given by few compact clusters including many patterns; in the case of symbolic methods, this is equivalent to few rules describing most

of the examples.

The CCE approach gives as a result a set of symbolic rules. We obtained few rules describing the class of "ill" patients. The class "non-ill" requires many more rules. This fact means that the ill subjects have common features that are not shared by non-ill subjects; in other words, we can say that these rules have caught the characterizing aspects of the data with respect to the Lyme problem.

With the AABP and Kohonen methods, we obtain a numeric result, that can be interpreted geometrically as clustering of points (each point is a pattern). The AABP approach led to the same conclusions of the CCE approach: unaffected patients lie within a very thick and concentrated cluster, whereas ill patients do not aggregate in a single group but appear shattered in a wider region. This evidences good class separation in homogenous clusters - a notable result by an unsupervised method. Moreover, the fact that a supervised (CCE) and an unsupervised method (AABP) lead to a similar conclusion about data structures is a very important result of our research, showing the congruence and correctness of current clinical representation. Fig. 3 displays a sample of a mono-dimensional distribution of middle-layer activation.

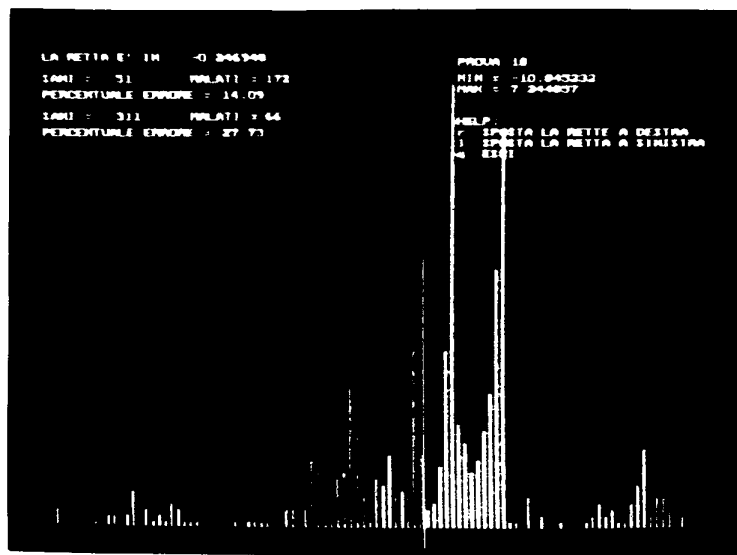


Fig. 3 Clustering with AABP

It can be observed that non supervised feed-forward networks have been used both for data compression and for clustering; in the first case we ask for an internal representation, while in the latter we are searching for a readable representation. As a consequence, it is not desirable to use more than one or two hidden units in the case of clustering, while in the case of compression we are free to choose the needed dimension for the hidden layer based on other considerations.

5: Future Lines of Research

The use of different neural methodologies to approach the same problem from different points of view has given as a result an effective cross-validation between the procedures, while at the same time confirming the correctness of the data chosen for the Lyme problem. Another cross-validation, using Kohonen's self organizing maps, is on plan.

A further step will be the application of the same methodologies to an augmented data set, added with new cases recorded by the Italian Working Group on Lyme disease. The Lyme diagnoses for these data will also be validated with the aid of recently introduced clinical methodologies, such as the PCR technique.

References

- [1] Barinaga M., *Furor at Lyme Disease Conference*, Science, Vol. 256, June 5, 1992
- [2] Bianchi G., Monteforte P., Rovetta G., Mancardi G.L., Crovato F., *Espressività della Borreliosi di Lyme in Liguria*, XV World Congress of Anatomic and Clinical Pathology, Abstract 52, Firenze, 1989
- [3] Kramer M.A., *Nonlinear Principal Component Analysis using Autoassociative Neural Networks*, AIChE Journal Vol. 37, n. 2, Feb. 1991
- [4] Moneta C., Parodi G.C., Rovetta S., Zunino R., *Automated diagnosis and disease characterization using neural network analysis*, 1992 IEEE SMC Conference Proceedings
- [5] Ridella S., Speroni, Trebino P., Zunino R., *Class-Entropy Minimization Methods for Domain Analysis and Rule Extraction*, Neural Computing and Application, in press, 1994
- [6] Rumelhart D.E., McClelland J.L. (eds.), *Parallel Distributed Processing. Explorations in the Microstructure of Cognition*, MIT Press, 1986
- [7] Weber K., Burgdorfer W. (eds.), *Aspects of Lyme Borreliosis*, Springer-Verlag, 1992