MEDICAL DIAGNOSES BY ARTIFICIAL INTELLIGENCE PROCESS

G. GRIGNOLIO

Institute of Mathematics
University of Pavia (Italy)

SUMMARY

A group of medical data collected in the Clinical Obstetrics Gynaecology Institute of Pavia University and the New-born Pathology Division at St. Matteo Hospital of Pavia has been worked out by an Artificial Intelligence process which enables to infer prediction formulae in spite of the large number of interrelated variables.

The aim of this work has been to set a mathematical formula which could prognosticate the probability of being alive 30 days from the birth, for a foetus whose weight is below 2.000 g.

The data which have been used concern 92 premature new-born babies who are less than 2.000 g weight.

We have effected several analyses covering a vast range of medical information which are supposed to be related to the possibility for the baby to live.

The information have been translated into binary variables (104 for each case) as follows: they have been considered on a real axis and one or more threshold has been chosen for each of them, so that the delimited areas were equally inhabited; it was moreover taken into account the significancy especially given to some thresholds from the medical point of view.

In order to study the problem it was first used a computer version ([3]) of the traditional statistic method of linear regression ([1]) which, as it is well known, hasn't been able to give any satisfactory result because of *the large number* of the variables which are *related in a very complex way* (the large number is in comparison with the number of the cases).

This fact depends both on the data drawn from the medical analyses (which are certainly related to one another) and on the binary classification; (note that if a value exceeds a threshold, then it necessarily exceeds also all the inferior ones).

Therefore an A.I. ("artificial intelligence") method of E. Gagliardo ([2]) has been used reaching in this way a simple formula which is suggested in order to effect diagnoses with a smaller probability of error.

This method can be schematically outlined as follows:

1) The 92 cases are divided into sets: the "training set" from which A.I. "learns" and the "test set" on which A.I. verifies the discriminative formula.

The training set is furthermore divided into two classes: the babies who die within the first month and who survive longer.

Then the mass-centers A and B both of the first and of the second class and the middle point C of the two mass-centers are calculated.

2) To every variable is given a "value" D which calculates the contribution of that variable to the discrimination between the two classes.

3) The formula which is then "learnt" by A.I. to forecast whether a child will get over the 30 days of life is the following:

$$\sum_{i=1}^{104} Z_i \cdot [X_i - C_i] + 0.210 \lesseqgtr 0$$

where:

$$Y_i = B_i - C_i$$

$$k = \frac{+1 \cdot}{\Sigma_i Y_i^2 D_i}$$

$$Z_i = k \cdot Y_i \cdot D_i$$

and X_i indicates the variables of the specific case to be diagnosed.

The diagnoses have proved to be correct on the 80% of the cases.

It is possible to obtain not only the given formula but also the values indicated by Z_i describing the positive influence (life) or the negative one (death) of the i-th variable on the result.

These values are at present object of researches in the obstetric field.

The formula which is learnt by A.I. has a vast possibility of application in the medical field when a small number of cases, though in presence of many measurements for each case, limits the significancy of the traditional statistics.

Many thanks for their kind collaboration to Prof. Vittorio Danesino and Dr. Secondo Guaschino from the Obstetric and Gynaecology Institute of Pavia University and to Prof. Giorgio Rondini from the New-born Pathology Division of St. Matteo Hospital of Pavia.

BIBLIOGRAPHY

1) Chatterjee S., Price B.: *Regression analysis by example.* John Wiley, 1977.
2) Gagliardo E.: *An example of a test for adaptive techniques in a medical investigation. Computers Biol. Med., 5,* 1975.
3) Gagliardo E.: *Versione informatica per elaboratore elettronico della analisi matematica.* F. Angeli, Milano, 1982.