
NEURAL NETWORK APPROACH PREDICTION OF THE TYPE OF A COURSE OF MULTIPLE SCLEROSIS BY THE CLINICAL CHARACTERISTICS OF ITS DEBUT

Inna Panchenko, Tetyana Shatovska

Keywords: Multilayer Perceptron, Multiple Sclerosis.

ACM Classification Keywords: F.2.1 Numerical Algorithms and Problems

Introduction

Methods of Intellectual Data Analysis are not obtained commonly in medicine because of small data warehouses with different length cases, combination of various data types (integer attributes, qualitative ones, scales of measurements, intervals, and binary parameters), the missed data abundant. A proportion of object quantity and researched parameters does not conform to principles of statistics that interpreted of the received results from its importance inconvenient.

The aim of our research is the Neural Network (NN) prognosis of the illness debut affiliation with the types of Multiple Sclerosis (MS) development.

Analysis of Multiple Sclerosis

Expert has sorted sampling under consideration into three MS classes: Primary Progressive Type (PPT), Relapsed-Remitted Type (RT) and Secondary Progressive Type (SPT). The last two ones intersect – some cases of RT transform to SPT, which means for researcher their only probabilistic separability [Lang, 1997].

Data warehouse includes temporal and clinical parameters of MS debut of 52 PPT patients, 46 RT patients and 68 SPT cases.

Characteristics of the MS debut are:

1. Age of debut.
2. Gender.
3. Risk factors before a debut: stress, infections, insolation, somatic diseases, operations, cranial traumas, spinal traumas, labors, abortions, others.
4. Harbingers: before debut, long-term ones.
5. Temps of development of a debut: hours, within a week, weeks, and months.
6. Character of a debut: monosyndromic, oligosyndromic, polysyndromic.
7. Syndromes of debut: retrobulbar [optic] neuritis, pyramidal, cerebellar, sensitive, brainstem, sphincters disorders, others.
8. Onset of symptoms order: simultaneous, fractional.
9. Duration of a debut: under 1 month, under 3 months, above 3 months.
10. Debut way out: under 2 wks, 2 wks – 1 month, above 1 month.
11. Severity of debut: mild, average, serious.

Age of the beginning of disease was integer value (yrs), data in all sub-groups of parameters 3-11 represented as binary.

According to the probabilistic separability of the MS development types the using of Probabilistic NN seems to be native, but it showed higher baseline error then Multilayer Perceptron (the average value for verification group was 0.3681), because of that the NN Multilayer Perceptron (MLP) was chosen [Grisby, 1994].

According to goal of our research, supervised learned two-layer architecture error back propagation algorithm MLP selected. The number of nodes in its hidden layer was determined as the half-number of input (forty clinical characteristics) and output (three types of MS) variables [Fayyad,1996].

Due to indexes of entropy ranging on the importance, only following sub-group parameters, which were influence to belonging to the MS type, used by the MLP for training:

1. Acute (under week) rate of MS development.
2. A short (under 1-month) debut.
3. Prolonged (above 1 month) debut regression.
4. Sex of patients.
5. Slowed (months) rate of MS development.
6. Rapid (under 2 wks) debut regression.
7. Moderate (2 wks – 1 month) debut regression.
8. Debut of average (less than 3 months) duration.
9. Stroke-like rate of debut development.
10. Presence of infectious background as a risk factor.
11. Fractional onset of attributes of MS manifestation.

Researched sample results (table 1):

Table 1 - Results of training, testing and verification of MLP

| | Training sample | | | Test sample | | | Verifying sample | | |
|-------------------|-----------------|-----|-----|-------------|-----|-----|------------------|-----|-----|
| | RT | PPT | SPT | RT | PPT | SPT | RT | PPT | SPT |
| Total of patients | 24 | 23 | 36 | 9 | 14 | 18 | 13 | 14 | 14 |
| Correctly | 10 | 20 | 31 | 4 | 11 | 15 | 6 | 13 | 12 |
| Incorrectly | 14 | 3 | 5 | 5 | 3 | 3 | 7 | 1 | 2 |
| Unknown | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| RT | 10 | 0 | 3 | 4 | 0 | 1 | 6 | 0 | 2 |
| PPT | 2 | 20 | 2 | 1 | 11 | 2 | 0 | 13 | 0 |
| SPT | 12 | 3 | 31 | 4 | 3 | 15 | 7 | 1 | 12 |

According to the table, difference of expert estimation and MLP training result was minimal at the analysis of progressive MS types. Therefore, from 23 training sample PPT cases, 20 examples have recognized correctly,

but three ones MLP has rated to SPT. From 14 cases of the test sample three ones also have classified as SPT and the result of verifying sample were 13 PPT cases and one SPT case.

When classifying the examples, which the expert has considered as SPT type, from 36 cases of the training sample two ones have defined as PPT, and three ones have to as RT. In the test sample of 14 cases, two ones have recognized as PPT, and one as RT. In the verifying sample of 14 cases, two ones defined to RT.

At RT, MLP could not distinguish the most of examples correctly. At the MLP training 12 of 24 patients have been referred to SPT, and two ones to PPT; in test sample 4 from 9 patients were classified as SPT, and one as PPT; the verifying sample has been divided as 6 patients with RT and 7 with SPT [Han, 2000].

It is significant, that divergences MLP with the expert in classification PPT concerned only SPT, i.e. essentially another progressive type of MS development whereas discrepancies between SPT and RT were opposite, and that is indirect evidence of an ambiguous belonging to the class - type of development.

Perhaps, classification of RT as SPT testifies to probable future transformation of RT to SPT.

Conclusion

Thus, based on application of a NN MLP, classification attributes of the MS debut were established. These attributes allow regarding progressive types of development, especially PPT, as stable clinical forms with good chance for forecasting whereas Relapsed-Remitted Type characterize instability of class belonging which means for the majority of patients ability of transformation to Secondary Progressive Type of disease.

Current problem seems to us to require further research.

Bibliography

[Lang, 1997] Lang E.W., Pitts L.H., Damron S.L., Rutledge R. Outcome after severe head injury: Analysis of prediction based upon comparison of neural network versus logistic regression analysis. *Neurological Research* 1997; 19. P. 274- 280.

[Grisby, 1994] Grisby J., Kooker R., Hershberger J. Simulated neural network to predict outcomes, cost and length of stay among orthopaedic rehabilitation patients. *Arch. Phys. Med. Rehabil.* 1994. - Vol. 75. P. 1077-1082.

[Fayyad,1996] Fayyad U. M., G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy. *Advances in Knowledge Discovery and Data Mining*. AAAI/MIT Press, 1996

[Han, 2000]J. Han and M. Kamber. *Data Mining: Concepts and Techniques*. Morgan Kaufmann, 2000.

Authors' Information

Inna Panchenko – Kharkiv National University of Radioelectronics, Department of Artificial Intellect, P.O.Box: Lenin avenue 14, Ukraine, 61166; e-mail: inna_pan_o@mail.ru

Tetyana Shatovska– Kharkiv National University of Radioelectronics, Computer Science department, P.O.Box: Kharkiv-116, Lenin av. 14, Ukraine; e-mail: tanita_uk@mail.ru