Applications of discriminant analysis methods in medical diagnostics

T. B. Martyniuk[a], L. M. Kupershtein[b], A. V. Medvid[a], A. V. Kozhemiako[a], W. Wojcik[b], O. Yuchshenko[c]
[a]Vinnytsia National Technical University, [b]Vinnytsia Financial Economical University;
[b]Lublin University of Technology, Nadbystrzycka 38A, 20-618 Lublin, Poland;
[c]Kazakh National Technical University after K. I. Satpaev, Satpaeva 22, Almaty, Kazakhstan

ABSTRACT

In this work the process of forming and processing the discriminant functions (DF) on a base of neural network classifier is researched. The research goal is the analysis of simulation educt of the process of the output vector forming appurtenances to certain class for the proposed neural network classifier.

Key words: discriminant functions, neural network, medical diagnostics.

1. INTRODUCTION

One of priority directions of the use of neural network technologies is medical diagnostics with such popular fields, as economics, business, telecommunications, Internet, security etc. Hereby, advisable for medical express diagnostics to use neural network classifier on the base of one- and multilayer perceptrons[1,2].

Into such classifier discriminant analysis is widely used, that allows to compartment object with the certain kit of the signs (of evidence) to one of well-known classes (diagnostic). Hereby classification is carried out by decision rules which represent by itself linear classifier functions (LCF) in the form of linear equations received on basis of the educative information[3-5].

2. APPLICATIONS OF DISCRIMINANT ANALYSIS METHODS

Classic neural network classifier consists of the one layer perseptron and maximum choice bloc – maximum detector[1,2]. As a result on the output of *m* linear neurons consisting of perseptron the *m* DF $g_i(\mathbf{x})$ are forming such look as:

$$g_i(\mathbf{x}) = \sum_{j=1}^{n} w_{ij} \cdot x_j, \quad i = \overline{1,m}, \tag{1}$$

where x_j – *j*-th the component of n-dimensional input vector **x**; w_{ij} – weight coefficient of *j*-th output of *i*-th neuron; *m* – the number of classes.

On the exit maximum-detector forms *m*-dimensional output vector **y**, in which unit value of *l*-th element y_l points to appurtenance of input pattern **x** to *l*-th class, that is decision rule has such look:

$$y_l = \{1 \mid \max \quad g_l(\mathbf{x}), \qquad l = \overline{1,m}\} \Rightarrow \mathbf{x} \in C_l, \tag{2}$$

where $C = \left\{C_1, ..., C_m\right\}$ – the classes set.

The third-layer neural network is used in the proposed variant of the classifier (fig. 1). Such classifier is contained *n* sensor neural elements of the input layer, *m* linear neural elements of the hidden layer, *m* binary neural elements of output the layer. In the classifier inputs supplied input *n*-elements vector **x** of signs, and on it output *m*-elements output vector **y** is formed, that point to the appurtenance input vector **x** of signs to the certain class.

Optical Fibers and Their Applications 2012, edited by Ryszard S. Romaniuk, Waldemar Wójcik,
Proc. of SPIE Vol. 8698, 86980G · © 2012 SPIE · CCC code: 0277-786/12/$18 · doi: 10.1117/12.2019733

The classifier is executed sequentially such operations as:
1) the multiplication of the elements of n-element input vector **x** of signs on corresponding weight coefficients $w_{ij}^{(2)}$ of neural elements hidden layer, that are formed in the learning process;

2) the forming of the sum S_i of weighted input signals $x_1,\dots,x_n$, that is corresponding DF on the output of the hidden layer neural elements in the form:

$$S_i = \sum_{j=1}^{n} w_{ij}^{(2)} x_j, i = 1,\dots,m, \tag{3}$$

3) the using for feedback between the input layer elements the weights $w_{ij}^{(3)}$ of lateral connection in the form:

$$w_{ij}^{(3)} = \begin{cases} 0, \mathit{if}\ i = j, \\ -\varepsilon \le \dfrac{1}{m}, \mathit{if}\ i \ne j, \end{cases} \quad i,j = 1,\dots,m, \tag{4}$$

4) the using of activation function $f^1(S_i)$ on the lateral outputs of output layer binary neutral elements in the form:

$$f^1(S_i) = \begin{cases} S_i, \mathit{if}\ S_i > 0, \\ 0, \mathit{if}\ S_i \le 0, \end{cases} \tag{5}$$

5) the using of activation function $f^2(S_i)$ on the normal outputs of output layer binary neutral elements in the form:

$$f^2(S_i) = \begin{cases} 1; & S_1 > 0 \\ 0; & S_1 \le 0 \end{cases} \tag{6}$$

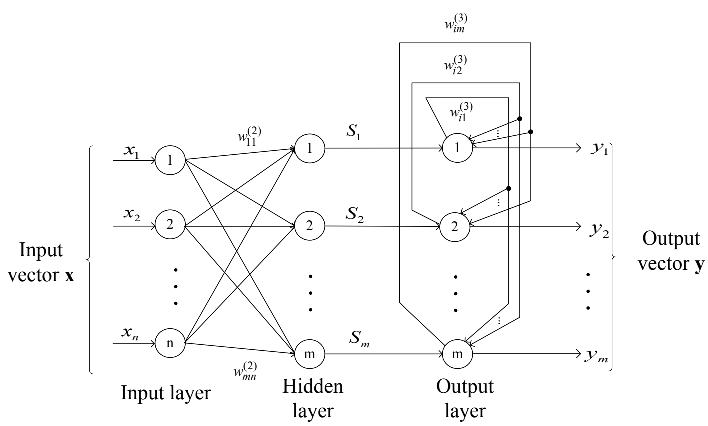

Figure 1. The structural organization of the neural network classifier

The iterative process of classifier functioning is finishing at the moment, when all binary neural elements of output layer, except one *l*-th element (winner with output signal not equal null), will go to the null state. Therefore, binary neural elements of output layer choose maximal discriminant function S_l (3) defining hereby *l*-th class, to which belongs input vector **x** of signs.

Peculiarity of proposed neural network classifier is that the connection weights $w_{ij}^{(2)}$ between sensory neural element outputs of input layer and linear neural element inputs of hidden layer are setting up in the learning process of classifier, and the lateral connection weights $w_{ij}^{(2)}$ of binary neural elements in output layer have constant values in the form (4).

In addition, in each binary neural elements of the output layer are used two types of activation functions, such as: the linear function (5) on the output of lateral connection and the Heaviside function (6) on the normal output.

Further the example of the classifier application in medical diagnostics is considered. In the learning process of the classifier the linear classification functions in a form of linear equation are formed, which are got by the methods of discriminant analysis on a base of learning information[3]. The learning information is formed for the results of the objects examination (patients) that are characterized by numerous signs (symptoms) and reliable fact of appurtenances to one of differential states (diagnosis). For each group of illnesses the generic linear classification function (LCF) is defined, which is contained in model all necessary symptoms[3]:

$$LCF_i = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n, \tag{7}$$

where LCF_i – linear classification function for i-th illnesses group, $i = 1, \ldots, m$; b_0 – constant; $b_1, b_2, \ldots, b_n$ – coefficients for the symptoms $x_1, x_2, \ldots, x_n$ respectively.

Eight symptoms $x_1, \ldots, x_8$ are selected for the diagnostics of three types of appendicitis (1 − gangrenous, 2 − phlegmonous, 3 − catarrhal) and another pathology of the alvus (tabl. 1)[3].

Table 1. Symptoms table coding.

Symptoms	Symptom name	Degrees of symptom expressiveness and their codes
x_1	Pains in the right spine area	1 – small; 2 - pronounced
x_2	Duration of pains in the right spine area	1 – more 2 days; 2 – 25-48 hour; 3 – 13-24 hour; 4 – to 12 hour
x_3	Pulse rate	1 – to 80; 2 – 81-100; 3 – more 100 ictus per min
x_4	Leukocytes	1 – to 8; 2 – 8-14; 3 – more $14 \cdot 10^9$ per liter
x_5	Tongue alterations	0 − not coated; 1 − coated
x_6	Blumberg's symptom	0 – no; 2 − pronounced
x_7	Rovsing's symptom	0 – no; 2 - pronounced
x_8	Protective muscle tension	0 – no; 2 - pronounced

For data of 103 medical report of patients with the three types of acute appendicitis and unconfirmed diagnosis are formed the matrix of learning information, in which in first column the diagnosis code is presented, what specifies, to what class patient belongs. In another eight columns of matrix 8 symptoms are specified. Some data of information array for classifier learning are presented in tabl. 2[3].

Table 2. The part of matrix learning information.

Diagnosis code	Symptoms							
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
1	2	3	3	2	1	2	0	2
1	2	3	1	2	1	2	2	0
2	2	3	1	2	1	2	2	2
2	2	4	1	2	0	2	2	0
3	2	4	2	2	1	2	0	2
3	2	3	2	2	1	2	2	2
4	1	2	1	1	0	0	0	0
4	1	1	2	1	0	0	2	0

Linear classification functions (LCF) for diagnostics of three types of appendicitis have such look[3]:

$$\begin{aligned}
LCF_1 &= -63,0 + 9,8x_1 + 3,6x_2 + 7,8x_3 + 5,2x_4 + 14,3x_6 + 11,8x_7 + 11,3x_8; \\
LCF_2 &= -57,4 + 8,3x_1 + 4,9x_2 + 6,2x_3 + 4,3x_4 + 13,5x_6 + 11,7x_7 + 10,6x_8; \\
LCF_3 &= -49,6 + 9,4x_1 + 4,7x_2 + 5,5x_3 + 3,0x_4 + 12,3x_6 + 12,0x_7 + 8,3x_8; \\
LCF_4 &= -23,0 + 6,3x_1 + 2,5x_2 + 5,3x_3 + 2,8x_4 + 7,8x_6 + 7,0x_7 + 5,8x_8.
\end{aligned} \tag{8}$$

The classification process in neural classifier is simulated in MathCAD 14 for the data from the tabl. 2.
For the example the simulation process of classification using LCF are adduced for the first illness group (the gangrenous appendicitis). There such values of input signs are choused from tabl. 2:

$$x_1 = 2,\ x_2 = 3,\ x_3 = 3,\ x_4 = 2,\ x_6 = 2,\ x_7 = 0,\ x_8 = 2. \tag{9}$$

The results of simulation in this case must denote on the first group of illnesses. The calculated LCF by the formula (8) have values:

$$LCF = \begin{pmatrix} 52{,}4 \\ 49{,}3 \\ 47 \\ 45{,}8 \end{pmatrix}. \tag{10}$$

The program is developed that simulate the algorithm of classification process by formulas (3)−(6), (8).
As a simulation result the vector is received:

$$S = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 4 \end{pmatrix}, \tag{11}$$

in which the "1" value of first element is pointed at the appurtenances of input vector **x** of signs to the first class. The fifth element of the vector S is pointed the number of cycles, for which maximal LCF was defined in classifier. The using of developed model was implemented on several sets of the vector **x** of signs for all four diagnostics classes. All results are coincided with data from tabl. 2 and proved convergence of the classification process by the formulas (3)-(6).

3. CONCLUSION

The classification process is simulated in neural network classifier, in which lateral connections are modified in output layer that allows paring the forming time of the output vector appurtenances considerably. The classifier functional principles are shown on the example of developed simulation model and confirmed the expected result of calculate by the classifier math model. The simulation result is received in the form of the vector, in which "1" value of l-th element points at the appurtenances of input vector **x** of signs to l-th class. The last element of the vector point at a number of cycles, for what define maximal LCF in classifier, what will admit to determine dependence time classification (the number of cycles) from dimension of input and output vectors of neural network classifier.

REFERENCES

[1] Haykin, S., [Neural Networks: A Comprehensive Foundation, Second Edition], Prentice Hall, Inc. (1999).

[2] Osowski, S., [Sieci Neuronowe do przetwarzania informacji], Oficyna Wydawnicza Politechniki Warszawskiej. Warszawa (2000).

[3] Юнкеров, В.И. [Математико-статистическая обработка данных медицинских исследований], СПб: ВМедА, (2002).

[4] Eishbeis, R. and Avery, R., [Discriminant Analysis and Classification Procedure : Theory and Applications], D.S. Heath and Company, New York (1972).

[5] Бернюков, А.К., "Распознавание биоэлектрических сигналов," Зарубежная радиоэлектроника 12, 47-51 (1996).