ACTA UNIVERSITATIS AGRICULTURAE ET SILVICULTURAE MENDELIANAE BRUNENSIS

Volume 62 79 Number 4, 2014

http://dx.doi.org/10.11118/actaun201462040769

NEURAL NETWORK FOR DETERMINING RISK RATE OF POST-HEART STROKE PATIENTS

Oldřich Trenz[1], Milan Sepši[2], Vladimír Konečný[1]

[1] Department of Informatics, Faculty of Business and Economics, Mendel University in Brno, 613 00 Brno, Czech Republic

[2] Department of Internal Cardiology Medicine – Institutions Shared with the Faculty Hospital in Brno – Institutions of Adult Age Medicine – Faculty of Medicine, Masaryk university in Brno, Jihlavská 20, 625 00 Brno, Czech Republic

Abstract

TRENZ OLDŘICH, SEPŠI MILAN, KONEČNÝ VLADIMÍR. 2014. Neural Network for Determining Risk Rate of Post-Heart Stroke Patients. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis,* 62(4): 769–776.

The ischemic heart disease presents an important health problem that affects a great part of the population and is the cause of one third of all deaths in the Czech Republic. The availability of data describing the patients' prognosis enables their further analysis, with the aim of lowering the patients' risk, by proposing optimum treatment. The main reason for creating the neural network model is not only to automate the process of establishing the risk rate of patients suffering from ischemic heart disease, but also to adapt it for practical use in clinical conditions. Our aim is to identify especially the specific group of risk-rate patients whose well-timed preventive care can improve the quality and prolong the length of their lives.

The aim of the paper is to propose a patient-parameter structure, using which we could create a suitable model based on a self-taught neural network. The emphasis is placed on identifying key descriptive parameters (in the form of a reduction of the available descriptive parameters) that are crucial for identifying the required patients, and simultaneously to achieve a portability of the model among individual clinical workplaces (availability of parameters).

Keywords: self-learning neural network, risk stratification, myocardial infarction

INTRODUCTION

For the target model creation, a self-teaching neural network was selected. It enables to determine – in a representative (teaching) number of patients – sets (clusters) of patients in various degrees of risk and their representatives (centres of gravity). By using of the rule of the tested patient's parameters proximity to the set representative, it is possible even in the calculation regime to determine the appurtenance of the tested patient to the appropriate risk set (Konečný, Trenz, 2009). For the classification itself, the model of the neural network is sufficient with the number of outputs equal to the number of risk sets; nevertheless, for the purpose of the description of the classified patient within the set, a planar depiction of the classification sets is used, the so-called Kohonen's map (Kohonen, 2001).

MATERIALS AND METHODS

Model Optimization

The initial model on which the possibility of neural networks exploitation was verified, contains twenty-six parameters (see Tab. I) characterizing the patient's ischemic illness (Konečný *et al.*, 2012). Neither the number of parameters nor their composition is suitable for practical usage, and so steps were taken to perform a detailed examination of their reduction based on:

- An expert opinion of a cardiologist,
- coefficients of parameters correlation,
- and the results of experiments.

The *primary* (Y) parameters were selected by a cardiologist (see Tab. I, recommendation column) that cannot be omitted in the model because they influence principally the risk assessment; *secondary* (N) parameters that will not be accepted in the model, and *uncertain* (Y/N) parameters, in which it is not possible, without an experiment, to determine unambiguously whether to leave them in the model or not.

The elimination of the parameter "Drugs" has been done mainly because of the fact that it is not proper to assess the degree of patients' risk conformably to the drugs prescribed to them; it should be done according to the principal measurable parameters whose dimensions are being affected by the drug's use.

The elimination of the parameters of the variability of the heart beat (HRV) from the model is supported by their mutual correlative dependence (Meloun, Militký, 2012), see Tab. II. With regard to the weakest correlation structure, the HRV Hfreq (21) and the HRV SDANN (24) have been kept in the primary parameters structure, i.e., one from the spectral analysis set and the other from the *time analysis set*.

The parameters of the pulse frequency (TF) during the heart stroke (IM), hypertension (HT), hyperlipoproteinemia (OA HLP) and diabetes (OA DM) did not markedly influence the classification function of the neural network and so have been removed from the group of the model base parameters (Konečný *et al.*, 2013).

In order to achieve the required classification into *low risk, medium risk* and *higher risk* patients groups and to avoid the classification according to the binary values of the *gender* parameter, this parameter has been ruled out of the S_{0M} base structure. The *gender* parameter caused that, during the self-teaching process with classification into three sub sets, there arose a separate sub set for the "female" gender and two sub sets for the "male" gender. This, however, does not mean that it is necessary to create a separate model for each gender.

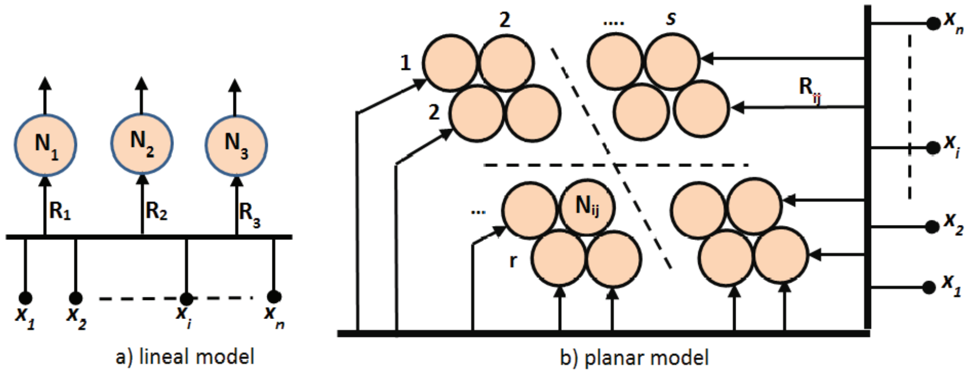
The parameters that are included in the S_{0M} structure form the *base structure*, i.e., a total of ten parameters altogether. This model contains all the primary as well as secondary parameters. The classification of any other variant of the model with a reduced parameters structure will be assessed according to its relationship to the base model. The $S_1 \div S_5$ parameter structures of the tested models are presented in Tab. I.

I: *Model Parameters*

No.	Parameter	Recommendation	S_0	S_1	S_2	S_3	S_4	S_5
1	Gender	Y	N	N	N	N	N	N
2	Age	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
3	ASA drugs	N	N	N	N	N	N	N
4	BB drugs	N	N	N	N	N	N	N
5	ACEI drugs	N	N	N	N	N	N	N
6	Statin drugs	N	N	N	N	N	N	N
7	Diuretic drugs	N	N	N	N	N	N	N
8	OA HLP	N	N	N	N	N	N	N
9	OA HT	N	N	N	N	N	N	N
10	OA DM	N	N	N	N	N	N	N
11	SKG revascular	Y	**Y**	**Y**	**Y**	N	N	N
12	SKG number of arteries	Y/N	**Y**	**Y**	**Y**	**Y**	**Y**	N
13	EF stratification	Y	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
14	QRS at IM	Y/N	**Y**	**Y**	**Y**	**Y**	N	N
15	TF at IM	N	N	N	N	N	N	N
16	DC	Y	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
17	AC	Y	**Y**	**Y**	N	N	N	N
18	Holter KES	Y/N	**Y**	N	N	N	N	N
19	HRV VLfreq	N	N	N	N	N	N	N
20	HRV Lfreq	N	N	N	N	N	N	N
21	HRV Hfreq	Y	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
22	HRV Mean NN	N	N	N	N	N	N	N
23	HRV SDNN	N	N	N	N	N	N	N
24	HRV SDANN	Y	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
25	HRV ASDNN	N	N	N	N	N	N	N
26	HRV RMSSD	N	N	N	N	N	N	N

II: *Correlation Coefficients*

PARAMETER	17	18	19	20	21	22	23	24	25	26
17	1.000	−0.450	−0.350	−0.352	−0.245	−0.157	−0.294	−0.241	−0.359	−0.183
18	−0.450	1.000	−0.066	0.015	−0.020	−0.213	−0.005	−0.005	−0.049	−0.066
19	−0.350	−0.066	**1.000**	**0.879**	**0.700**	**0.613**	**0.758**	**0.672**	**0.971**	**0.645**
20	−0.352	0.015	**0.879**	**1.000**	**0.786**	0.395	**0.654**	**0.582**	**0.894**	**0.632**
21	−0.245	−0.020	**0.700**	**0.786**	**1.000**	0.460	**0.554**	0.481	**0.777**	**0.877**
22	−0.157	−0.213	**0.613**	0.395	0.460	**1.000**	**0.548**	**0.485**	**0.630**	**0.547**
23	−0.294	−0.005	**0.758**	**0.654**	**0.554**	**0.548**	**1.000**	**0.988**	**0.767**	**0.570**
24	−0.241	−0.005	**0.672**	**0.582**	0.481	0.485	**0.988**	**1.000**	**0.681**	**0.502**
25	−0.359	−0.049	**0.971**	**0.894**	**0.777**	**0.630**	**0.767**	**0.681**	**1.000**	**0.737**
26	−0.183	−0.066	**0.645**	**0.632**	**0.877**	**0.547**	**0.570**	**0.502**	**0.737**	**1.000**



1: *Model Types*

Model Types

With regard to medical practice, a classification of patients into three sub sets has been selected: *low risk, medium risk* and *higher risk* patients, while exploiting the self-teaching neural network. There is no pre-established teaching result for this network; nevertheless, it enables us to determine the existing clusters (sub sets) in the chosen set of input objects (the teaching set), characterized by quantifiable parameters. Two variants of the network are used:

- In a one-dimensional setting (lineal model) of 3 × 1 outputs, see Fig. 1a.
- In a two-dimensional setting (plane model) of $r \times s$ outputs, see Fig. 1b.

Furthermore, the abovementioned relations are valid for both versions, bearing in mind that, in the case of the version with a single-dimensional output setting, the line-index is equal to one or is, for simplicity's sake, left out.

The neural network inputs form vector coordinates $X = (x_1, x_2, \ldots, x_n)$ of the classified objects. The individual vectors consist of measurable parameters of patients, the structure of which is established – for each model – by the accepted parameters, according to Tab. I.

Every input vector X activates only that output neuron, whose R_{ij} vector corresponds to:

$$|R_{ij} - X| \leq |R_{kl} - X|, \forall\, k, l; k \neq i, l \neq j. \quad (1)$$

This means that the taught neural network performs a classification of input vectors (objects) into sub sets corresponding to the neural network's outputs. In the case of networks on Fig. 1a it will be a classification into three sub sets, and in the case of the network on Fig. 1b, into $r \times s$ sub sets.

We may perform the display of classification sub sets represented by output neurons and their color differentiation in a planar model. After finishing the self-teaching process of the lineal model, every output neuron with the R_{ij} vector will receive a colour corresponding to the classification set represented by the R_k vector of the lineal model, when we can say that:

$$|R_k - R_{ij}| \leq |R_l - R_{ij}|, \forall\, l, l \in \{1, 2, 3\}. \quad (2)$$

Then we can say, for any input vector X of the planar model, that:

$$(X \in M_{ij}) \wedge (M_{ij} \subset M_k) \Rightarrow X \in M_k, \quad (3)$$

where M_k is the set represented by the R_k vector and M_{ij} is the set represented by the R_{ij} vector.

The self-teaching process of the neural network itself consists of correcting the R_{ij} (or $R_k = R_{1k}$) vector of that output neuron which, with the input vector, abides by condition (1). The correction is done according to the relation:

$$R^*_{ij} = R_{ij} + \alpha(\tau)(X - R_{ij}). \quad (4)$$

The teaching coefficient $\alpha(\tau) < 1$ and changes with the increasing number of teaching epochs τ. Simultaneously with correcting the vector of the victor neuron, a correction of the vectors of its neighboring neurons is done, according to:

$$R^*_{ij} + \alpha(\tau)\beta(\tau, d)(X - R_{ij}). \quad (5)$$

The teaching coefficient $\beta(\tau, d) < 1$ changes with the teaching epoch, as well as with the neighbour's distance from the victor neuron. The teaching-group objects' selection is done randomly. The teaching process is discussed in greater detail in (Konečný, Trenz, 2009) and (Konečný *et al.*, 2013).

The classification sets representatives' vectors can be established (Konečný, Trenz, 2009) for every parameters structure:

- *self-teaching of the neural network with the number of outputs equal to the number of classification groups* (in the case of classifying patients into three sub sets, see neural network on Fig. 1a),
- or as *a centre of gravity of classification-sets' objects*, explicitly established by an expert.

In the case that the set specification is established vis á vis the neural network's ability to detect object clusters, and not by an expert, we consider the model classification on Fig. 1a with the S_0 structure to be the exemplary (or rather, expert) one.

Although every model with the S_i structure (see Tab. I) will employ, for establishing representative vectors (centre of gravity), classification sets of the initial S_0 model, as a consequence of the elimination of checked parameters, the representatives' vectors, the mutual position of the representatives and of the classified objects (and consequently also the classification) will differ. An accepted error of the model with a reduced structure will be less than 5% of objects that have been classified otherwise.

RESULTS

Input Data Handling

For the purpose of individual parameters weight adjustment it is suitable, as step number one, to perform the parameter standardization so that the standard deviation σ_i is equal to one, and the mid-value $\bar{x}_i$ of each parameter is equal to zero (Meloun, Militký, 2012).

So that the classification sets do not undergo an unwanted deformation, it is suitable to remove, from the teaching file, objects that are extremely distant from the centre of gravity of all the teaching file objects set. The appearance of such objects may be caused by data measurement, subtraction, or recording errors. Those objects, for which the relation

$$|X_k - R_0| > 1.25\sigma_0, \quad (6)$$

is valid (where X_k is the k-th object's (the patient's) coordinate vector of the teaching file, R_0 is the vector of the centre of gravity of the input objects set, and σ_0 is the standard deviation of the objects' distance from the centre of gravity). Because upon performing the standardization, $\bar{x}_i = 0$ holds for all the parameters, the centre of gravity vector will be equal to the zero vector.

Establishing the Sample Classification

The basis for self-taught neural network object classification is the rule of proximity, according to which two equal objects (with a zero distance) have the same quantitative parameters, and two proximate objects also possess cognate parameters, and thus their characteristics should also be similar or proximate.

We may easily verify that the sum of distances $d_i(X, R)$ of objects with an X position vector from an object with an R position vector is the smallest when the R vector represents the set's centre of gravity. In the self-teaching process, we perform an adjustment of the output vector R_k that is closest. The result of the progressive R_1, R_2 a R_3 (Fig. 1a) vectors adjustment are the position vectors of the classification sub-sets $M_1(R_1)$, $M_2(R_2)$ and $M_3(R_3)$.

Due to the fact that every input vector X activates only the single R_k vector output for which the relation (1) is valid, we may easily determine the elements corresponding to the classification sub-set $M_k(R_k)$. In the case of a model with the S_0 parameters structure, the $M^0_k(R^0_k)$ classification sets are considered to be exemplary. Every other classification's correctness evaluation is done by contrasting it with the exemplary classification.

Self-Teaching of the Planar Model

The teaching process is performed in the same way as in the case of the lineal model, the only difference being a greater number of outputs. The planar model's outputs classification is performed according to the distance of vectors of R_{ij}-outputs of the planar model from the vectors of the row model R_k – see (2). Nevertheless, a certain classification error rate – in contrast with the row model – may be expected as a consequence of the imprecise teaching of the network, and usually in the case of borderline objects of the sets.

Upon reaching a coherent coloring of the output sets in the planar model, we have used the methodology of data adjustment of the teaching file that is described in detail in the publication (Konečný *et al.*, 2013).

Model Change

Each examined model requires an adjustment of the teaching file of the neural network according to the structure of S_i parameters and establishing the vectors of R_k (centre of gravity) representatives of classification sets. Both these operations represent a reduction of X_i input vector and R^0_k representative

III: *Data Sample of the Teaching File*

Identifier	Age	EF Stratification			DC	HRAV Hfreq	HRV SDNN
	2	13			16	21	24
m1	71	49			5.2	99	160
m3	66	37			3.85	41	89
m4	67	50			5.71	44	75
m6	75	27			4.63	52	85
m8	54	45			4.94	26	85
m9	59	53	…	…	9.59	287	112
…							
m85	53	45			7.40	233	127
m87	59	24			–4.46	57	59
m89	46	70			8.41	106	96
m90	61	45			9.01	167	120

IV: *Experiment Results*

		Gender: Male						Gender: Female						
Model Parameter Structure →		S_1	S_2	S_3	S_4	S_5	$\lvert M_i \rvert$*	S_0	S_1	S_2	S_3	S_4	S_5	$\lvert M_i \rvert$*
3 × 1 outputs	M_1	0/0	0/0	0/0	4/9.52	13/31	39	0/0	0/0	0/0	2/16.7	3/25.0	6/50.0	12
	M_2	0/0	0/0	2/4.6	5/11.63	16/37.2	45	0/0	0/0	0/0	2/11.1	4/22.2	6/33.3	18
	M_3	0/0	0/0	2/9.5	1/4.76	3/14.3	22	0/0	0/0	0/0	0/0	3/60.0	2/40.0	5
Total		0/0	0/0	2/1.9	**5/4.72**	16/14.3	106	0/0	0/0	0/0	**2/5.7**	5/14.3	7/20	35
r × *s* outputs	M_1	1/2.4	2/4.8	1/2.4	4/9.52	13/31	39	0/0	0/0	1/8.3	3/25.0	4/33.3	7/58.3	12
	M_2	2/4.8	3/7.1	4/9.5	6/14.29	18/42.9	45	0/0	0/0	1/5.6	3/16.7	5/27.8	8/44.4	18
	M_3	1/4.6	1/4.6	3/13.6	2/9.09	5/22.7	22	0/0	0/0	0/0	0/0	3/60.0	3/60.0	5
Total		2/1.9	3/2.8	4/3.8	6/5.7	18/17	106	0/0	0/0	1/2.9	3/8.6	6/17.1	9/25.7	35

$\lvert M_i \rvert$* – the set mass

vector parameters of the S_0 representative structure, according to the new S_i structure.

The logical supposition that the model with a base structure of S_0 parameters will be more precise in the classification than the simplified model has lead to the fact that the basis for establishing the R_k output vector (centre of gravity of the $M_k(R_k)$ set of the simplified model) is the $M_k^0(R_k^0)$ classification set. The coordinates of the w_i vector of the R_k can then be established according to the relation:

$$w_i = \frac{1}{N_k^0} \sum_{j=1}^{N_k^0} x_{ji} \,, \tag{7}$$

where N_k^0 is the number of elements of the $M_k^0(R_k^0)$ set and x_{ji} is the i-th coordinate of the vector of the j-th element of the $M_k^0(R_k^0)$ classification set.

This approach will enable us to verify the hypothesis of whether classifying using a model in which the R_k vectors are derived from the R_k^0 vectors of the base model will be most precise.

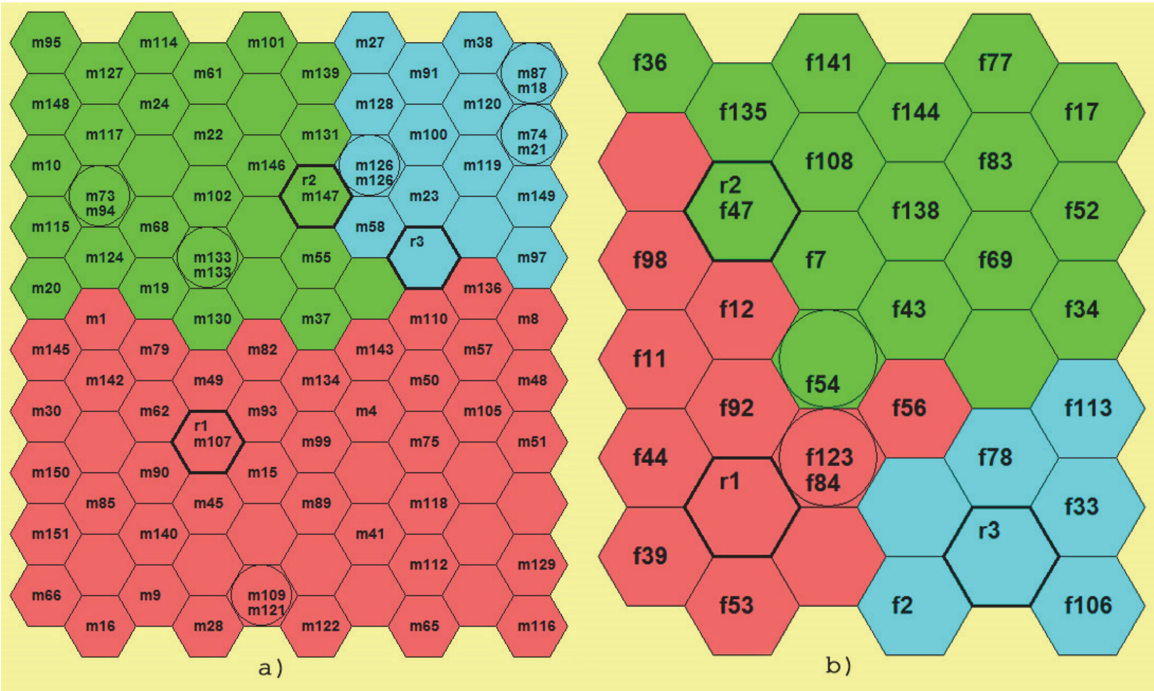
The second option for establishing the R_k vectors is by means of the neural network's self-teaching, in Fig. 1a, independently for each S_i parameters structure, although the error rate – when contrasted with the base structure – was worse.

Suitable Model Selection

The basis for creating models and performing the experiments was the data of patients undergoing treatment at the Clinic of Cardiology of the Brno Faculty Hospital (Sepši, 2008). In the „Male“ model we worked with 106 patients, and in the „Female“ model, with 35 patients. The data sample of the neural network teaching file for the S_5 structure model is listed in Tab. III.

The test-model experiment results – individually for both genders – are listed in Tab. IV. Individual cells of the chart contain data about the number of erroneously classified objects and their percentage share in relation to the overall number of elements of the $M_k^0(R_k^0)$ set.

What was decisive for assessing the classification was the model with 3 × 1 outputs, because in this model we are deciding about the classification according to the distance of the classified object directly from the set's centre of gravity, whereas in the planar model we decide about the output classification according to the R_{ij} output vectors distance (established in the self-teaching process) from the R_k representatives of the classification groups. The discrepancies in the classification



2: *Displaying the Model's Classification Groups for the Male Gender*

V: *Model Mortality Classification*

Model Parameter Structure →		Gender: Male						Gender: Female					
		S_0	S_1	S_2	S_3	S_4	S_5	S_0	S_1	S_2	S_3	S_4	S_5
2 × 1 outputs	M_1	4/0.1	4/0.1	4/0.1	4/0.1	4/0.09	5/0.1	2/0.17	2/0.17	2/0.17	2/0.17	2/0.18	3/0.21
	M_2	11/0.3	11/0.3	11/0.3	12/0.3	12/0.30	10/0.2	5/0.28	5/0.28	5/0.28	5/0.28	4/0.22	4/0.25
	M_3	11/0.5	11/0.5	11/0.5	10/0.5	10/0.50	11/0.5	4/0.80	4/0.8	4/0.8	4/0.8	5/0.83	4/0.8
r × *s* outputs	M_1	4/0.1	4/0.1	4/0.1	4/0.1	4/0.09	5/0.1	2/0.17	2/0.17	1/0.09	1/0.09	1/0.11	2/0.15
	M_2	10/0.2	11/0.3	11/0.3	12/0.3	12/0.30	10/0.2	5/0.28	5/0.28	6/0.32	6/0.32	5/0.26	6/0.33
	M_3	12/0.6	11/0.5	11/0.5	10/0.5	10/0.50	11/0.5	4/0.80	4/0.8	4/0.8	4/0.8	5/0.83	3/0.75

of these two models have occurred most frequently on the borders of sets.

From the viewpoint of the parameters number (a total of six) and the achieved precision of the classification (less than 5% of erroneously classified objects out of the total number), the suitable model – for the *male* gender – seems to be a model with the parameter structure S_4 and for the *female* gender – the model with the structure S_3, which slightly surpasses the requirements for classification precision. Further model simplifications are not necessary. They would cause an unwanted increase in the number of erroneously classified objects. The S_4 structure model, as we may see in Tab. I, contains merely *the primary elements* (Y) which have been defined by an expert with certainty, except for the AC parameter, which has been left out based on the results of the performed experiments.

Displaying the teaching file sets by means of the model on Fig. 1b. with the S_4 structure – gender:male is shown on Fig. 2a, while the S_3 – gender:female structure is shown on Fig. 2b. Some outputs are multiple (employed by various objects) and thus we cannot display all the elements on the image. The strongly traced outputs with r1, r2, and r3 identifiers represent objects with classification sets representatives (centre of gravity). The tested objects appear in the circle with the identifier at the bottom.

2.3 Semantics of the Classification Groups

The classification-sets patients' risk rate is derived from the mortality indicator *m*, which has been established for every classification group of the $S_0 \div S_5$ models according to the relation:

$$m = \frac{N_{zem}}{N_{zem} + N_{ziv}}\,, \tag{8}$$

where N_{zem} is the number of deceased patients of the class, and N_{ziv} is the number of living patients. The overall number of healthy/deceased patients classified in the $M(R_1)$, $M(R_2)$ and $M(R_3)$ and sets is 106/26 in the males model and 35/11 in the females model. The results of the mortality are listed in Tab. V.

VI: *Patients' Age Structure*

	Men			Women		
	M_1	M_2	M_3	M_1	M_2	M_3
$\|M_i\|$	53	35	53	12	18	5
Age > = 85	1	10	14	0	1	4
65 < Age < 85	39	11	39	9	8	0
Age < = 65	13	14	0	3	9	1
Average age	60.89	71.51	72.21	74	66.94	82.00

The individual chart cells contain the N_{zem}/m data. We may gather from Tab. V that the differences in the deceased patients' classification by different models are only two patients in the case of the 3 × 1 outputs model, and one patient in the $r \times s$ outputs model.

According to the calculated mortality rate we may judge that Class 1 represents patients with lesser risk, Class 22 patients with a greater risk, and Class 3 patients with the greatest risk.

The semantics of the classification sets cannot be established by means of the age structure of the patients of these sets (see Tab. VI). To simplify, we may suppose that patients older than eighty will have a *greater risk rate,* patients younger than 65 years *a low risk rate,* and the other patients will have a *mid-value risk rate*. Nevertheless, comparable numbers of patients of a similar age interval and of different classes do exclude it as a classification rule.

A Model Accepting the Parameters' Weight

While practically assessing the objects' characteristics, all the parameters are not usually accepted as having the same importance. The levels of importance are usually expressed in a suitable value interval and by the relation:

$$w_i = \frac{v_i}{\sum_i v_i} N \,, \tag{9}$$

where N is the number of parameters, the degrees of importance are transferred to normalized weights. The weights are used as standard deviations of the corresponding parameters. In the case that all the parameters have an equal importance, all the weights, and thus also all the standard deviations, are the same and are equal to one.

In the phase of parameter selection – with the aim to minimize their number – we have not used the option of creating a model with parameter weights. We consider these in the case of a further improvement of its function; for instance, if we come to the conclusion that we should use the ($w_i = 0$) parameter with the $w_i > 0$ mass.

DISCUSSION

For the creation of neural models we have used the data of 141 patients. 106 for the teaching group „Male" and 35 for the teaching group „Female". It is evident that the testing model will necessarily have to be refined by means of a continuous filling up of the neural network's teaching set (Chapelle *et al.*, 2006).

By eliminating the parameters for the final model structures, the precision of classification, as opposed to the initial S_0 structure, has worsened down to the accepted required value, i.e., a maximum of five percent of incorrectly classified patients.

By means of a reduction of parameters, we have achieved the required goal, i.e., to simplify to the maximum, and thus also to render accessible, the practical usability of the model (Hinton, 2006).

With the exception of specific software used while verifying the models and the planar display of the classification sets (or rather, the tested patients) we have verified a simplified application in the chart processor MS Excel, without the planar display (Kohonen's map). This follows up on the exemplary neural network or k-means classification (Škorpil, Šťastný, 2006; Trenz, Konečný, 2010).

SUMMARY

The present paper describes the issue of creating a decision-making model for the classification of patients' health. For building the model we have used the principle of self-organized Kohonen's networks, including the option of displaying the classified sets in the form of an output planar depiction (Kohonen's Map). A set of 141 patients with 26 descriptive parameters served as the input group of data to be verified.

Due to the large number of parameters, which would prove unsuitable for a global implementation of the decision-making model, we have set out to reduce them, based on the expert opinion of a cardiologist, their correlation, the results of performed experiments, and all this in the effort to reduce their number to the accepted level enabling us to perform our own decision-making at a random cardiologist workplace.

Based on experiments and upon consulting a cardiologist we have acceded to classifying patients with the help of the created decision-making model into three sets: *low risk rate, mid-value risk rate and high risk rate* patients. This differentiation enables us to apply preventive healing with the aim of reducing mainly the occurrence of sudden heart stroke death. A necessary condition of implementing the model is the initial data standardization, in order to eliminate the individual parameters' influence on the result. New and targeted use of weights is possible upon consulting an expert; it was not included in the abovementioned model.

For the purposes of further research, we have given priority to the S_4 model with six parameters identifying male-patients' health, and the S_3 model with seven parameters identifying female-patients' health. This parameter version enables us to identify the required patients' state with an acceptable error, and it is, for its simplicity, vis á vis the number of parameters and medical practice, readily usable.

REFERENCES

CHAPELLE, O., SCHÖLKOPF, B. and ZIEN, A. 2006. *Semi-supervised learning*. Cambridge, Mass.: MIT Press.

HINTON, G. E. 2006. Reducing the Dimensionality of Data with Neural Networks. *Science*. 313(5786): 504–507.

KOHONEN, T. 2001. *Self-organizing maps*. 3rd ed. Berlin: Springer-Verlag.

KONEČNÝ, V., SEPŠI, M. and TRENZ, O. 2012. Analysis of evaluation problems of the risk situation of patients suffering from ischemic heart disease. *Acta Univ. Agric. Silvic. Mendelianae Brun.*, 60(2): 125–134.

KONEČNÝ, V., TRENZ, O. 2009. Decision support with artificial intelligence. *Folia Univ. Agric. et Silvic. Mendelianae Brun.*, 2(8).

KONEČNÝ, V., TRENZ, O. and SEPŠI, M. 2013. Data Adjustment for the Purposes of Self-Teaching of the Neural Network, and its Application for the Model-Reduction of Classification of Patients Suffering from the Ischemic Heart Disease. *Acta Univ. Agric. Silvic. Mendelianae Brun.*, 61(2): 377–384.

MELOUN, M., MILITKÝ, J. 2012. *Kompendium statistického zpracování dat*. Vyd. 3. Praha: Karolinum.

SEPŠI, M. 2008. *Srovnání stratifikačních postupů k určení rizika náhlé srdeční smrti u nemocných po infarktu myokardu*. Ph.D. práce. Brno: Lékařská fakulta, Masarykova univerzita.

ŠKORPIL, V., ŠŤASTNÝ, J. 2006. Back-Propagation and K-Means Algorithms Comparison. In: *2006 8th International Conference on SIGNAL PROCESSING Proceedings*. Guilin, China: IEEE Press, 1871–1874.

TRENZ, O., KONEČNÝ, V. 2010. Comparison of the applicability of neural networks and cluster classification methods on the example company's financial situation. *Acta Univ. Agric. Silvic. Mendelianae Brun.*, 58(6): 579–585.

Contact information

Oldřich Trenz: oldrich.trenz@mendelu.cz
Vladimír Konečný: Vladi-mir.konecny@mendelu.cz
Milan Sepši: msepsi@fnbrno.cz