neural network,
expert systems, heart diseases

Arkadiusz POPA*, Sebastian WOJCZYK, Marcin LIZIS

SELF OPTIMIZING NEURAL NETWORK – AS EXPERT SYSTEM IN MEDICAL HEART ATTACK.

with the supervision of Andrzej Nowakowski

The main aim of Self Optimizing Neural Network (SONN), which are presented in this paper, is construction of expert system on the basis of analysis of medical information about group of patients. The expert system is built on the basis of neural network, and the main task of this system is to expect future patient health, based on information about the patient. Such a system can give the doctors a hint about that what can be happen with patient. And what is more important – the SONN construction process is very flexibly and adapts topology and all weights to training data. This is undoubtedly a great advantage of this type of neural network. Moreover the construction process is quite simple. The network topology and all connections between neurons can be easy implemented and kept in such a way, which allows to create very efficient expert system.

In this paper we describe the process of construction of neural network which is based on one-shot analysis of learning patterns. On the basis of appropriate computation the SONN topology is built. The construction process can be repeated on the larger group of patients. In this way the expert system (based on SONN) will be better and better.

1. TRAINING PATTERNS

In this paper the training data (training patterns) are patients. Each patient consist of set of medical data (which describe) which will be called „features". These features describe e.g. sex, age, Rrs, HR and other patient data. Choice of the best discriminative features is essential in SONN, because these features are basis to the network topology building.

In this work 17 the most important features (medical data about patient) were chosen. Then each of features was divided into intervals. Such a division was necessary and enabled data transformation from continuous to discrete form, e.g. some medical feature (called HR) was divided into IV intervals, respectively:

I) 0 – 45, II) 46 – 100, III) 101 – 150, IV) 151 – infinity.

In such a way each of 17 features was divided. The way of divided each feature is describe in section „SONN learning data transformation".

2. SONN LEARNING DATA TRANSFORMATION

The SONN construction process works only if learning data consist 1, -1 an 0. It mean that feature exists, not exists or we do not know if exists in learning pattern, respectively. In other words: if value of medical data is included in fixed interval, then the feature has value 1, and other intervals have value -1. If we have no information about value of medical data, all the features have value 0.

It is illustrated in following example. In the table is shown how the information about age and HR were transformed from continuous to discrete form (-1,1,0)

Table 1. Example of converting two medical features from continuous to discrete values.

Medical feature	Intervals	Patient A		Patient B	
		Medical value	Discrete value	Medical value	Discrete value
Age	0 – 63	55	1	lack of data	0
	64 – inf.		-1		0
HR	0 – 45	105	-1	47	-1
	46 – 100		-1		1
	101 – 150		1		-1
	151 – inf.		-1		-1

In the table is show transformation medical features value to discrete value according to set intervals. Intervals for particular medical features are shown below:

* University of Łódź, Faculty of Mathematics and Informatics, ul. Banacha 22, 90-238 Łódź, Poland

Table 2. Set intervals for medical features

Medical feature	**Intervals**	**Medical feature**	**Intervals**	**Medical feature**	**Intervals**	**Medical feature**	**Intervals**
1) Sex	Female	6) CKMB max	0 – 200	12) GO	0 - 75	17) EF%	0 - 30
	Male		201 – inf		76 - 99		31 - 45
2) Age	0 – 63	7) front	0 or 1		100		46 - inf
	64 – inf	8) front-bottom	0 or 1	13) GPZ	0 - 75		
					76 - 99		
					100		
3) RRS	0 – 90	9) bottom	0 or 1				
	91 – inf	10) PLTW	0 – 50				
4) HR	0 – 45		51 – 75	14)critical changes	0 or 1		
	46 – 100		76 – 99	15) TIMI start	0		
	101 – 150		100		1		
	151 – inf	11) GPZ	0 -75		2, 3		
5) CKMB	0 – 27		76 – 99	16) TIMI end	0, 1, 2		
					3		
	28 – 81		100				
	82 – inf						

The given features, which characterize particular patterns, identify output classes.
It was established 5 distinct output classes. (It means that: Patient with impact or Death was classified to A output class.
Five output classes:
A) death
B) re-impact
C) impact
D) revascularization
Z) live

3. COMPUTATION OF DISCRIMINATION COEFFICIENT

In this type of neural network the discrimination coefficients are computed only once. It means that for each pattern and each feature we make computation only once. The value of this coefficient described if the feature can be distinguished from the same feature in the other patterns. On the basis of this computation feature with maximal value of discrimination coefficient are chosen. It means that we choose only these features which discriminate the pattern the best. It ensures univocal classification patterns to properly class. The discrimination coefficient is computed according to mathematical formula.

$$\forall_{m\in\{1,\dots,M\}}\forall_{u^n\in C^m}\forall_{n\in\{1,\dots,Q\}}\forall_{k\in\{1,\dots,K\}}\begin{cases} d_k^{n+}=\dfrac{\hat{P}_k^m}{(M-1)Q^m}\displaystyle\sum_{\substack{h=1\\h\neq m}}^{M}\left(1-\dfrac{\hat{P}_h^h}{Q^h}\right) & \text{if} \quad x_k^n\neq 0\\ d_k^{n-}=\dfrac{\hat{N}_k^m}{(M-1)Q^m}\displaystyle\sum_{\substack{h=1\\h\neq m}}^{M}\left(1-\dfrac{\hat{N}_h^h}{Q^h}\right) & \text{if} \quad y_k^n\neq 0\end{cases}$$

where:

Q - number of patterns in classes
Q^m - number of patterns in class m
M - number of output classes
K - number of features
d_k^{n+} - discrimination coefficient for positive value feature
d_k^{n-} - discrimination coefficient for negative value feature
$\hat{P}_k^m$ - number of positive value features on feature k and class m in view of all training set

$\hat{N}_k^m$ - number of negative value features on feature k and class m in view of all training set

4. NETWORK TOPOLOGY

After computation of all discrimination coefficients we start with the construction of network topology. The algorithm can be described as follows:

1. Computation of discrimination coefficient for each feature and all training patterns
2. Maximal discrimination value for all features of each training pattern is chosen. If the pattern has more than one feature with maximal discrimination value, then the first is chosen. The chosen feature is the first network layer.
3. Each training pattern must be univocally represented by chosen features. If the feature which was chosen (or set of chosen features) is identically for several patterns (which belong to different classes), then the next feature with maximal discrimination coefficient is chosen (from features which we have not used so far). This is made for all patterns (We do not have to compute discrimination coefficient once again).
4. The step 3 is repeated until each of output classes be univocal classified by set of features with maximal discrimination coefficient. The set must be unique in view of all classes. Each consecutive iteration causes that new network layer is added.

Proper choice of number of SONN network layer is very important. The more properly choice guarantee best results during learning process and best result in network testing as well. But SONN network cannot be build by maximal number of layers. In that way neural network is going to remember pattern exactly as it is and is not immunity on testing troubles. The best results are when 75% of features used (SONN network use 75% of layers).

5. RESEARCHES

In order to build and learn the network we used 804 patterns. One pattern represents one patient. Each pattern consists of 40 features which characterize this pattern. Particular feature in pattern denotes result of medical checkup. The features must be discrete, which is necessary in view of neural network.

From group of 804 patterns it was chosen 665 random patterns, which will be used to building the neural network.

Remaining 139 patterns will be used as testing patterns. Knowing „correct answers" we would evaluate network answers for these testing patterns.

6. RESULTS OF SONN

From group of 804 patients, 665 were used to network building. Remaining 139 patients were used to network testing. The results of classification are:

I. In first case the network was built using 75% of features. (i.e. 30 features) Univocally classified: 116 patterns (83,45%)
   Equivocally classified: 3 patterns (2,16%)
   not classified: 20 (14,39 %)
II. Reducing the number of features from 30 to 20, we obtained the following result:
   Univocally classified: 109 patterns (78,42%)
   Equivocally classified: 7 patterns (5,04%)
   Not classified: 23 (16,55 %)
III. Further reduction of number of network layers to 15 gave the following result:
   Univocally classified: 91 patterns (65,47%)
   Equivocally classified: 21 patterns (15,11%)
   Not classified: 27 (19,42 %)

Remark 1. Reducing the number of features, which were use in network building, the number of univocally classified testing patterns is also reducing. The best results was achieved if the network was built using 75% of features. In this case increase number of patterns which are classified equivocally. But it may be a consequence of the fact, that the same features composition may suit to more than one training patterns

It was also built the network using only 296 learning patterns. The network was tested with 139 testing patterns.

I. Univocally classified: 105 patterns (75,54%)
   Equivocally classified: 3 patterns (2,16%)
   Not classified: 31 (22,3 %)
   The network was built using 75% of features (i.e. 30 features)

Remark 2. If construction of SONN use not enough learning patterns, the network could not classified all testing patterns.

Simple example of neural network building process:

It was chosen 20 random patterns. Each pattern (patient) is described by 4 medical features: sex, age, Rrs, HR. According to previous notes and division to intervals we obtain 10 features with discrete value. Selected medical features, including division and data transformation are shown in following table.

Table 3. Example of converting medical continuous to discrete features for chosen patient.

Feature description	Medical intervals	Discrete value	Patient A	
			Ex. feature	Discrete
sex	K	1 or -1 or 0	K	1
	M	1 or -1 or 0		-1
age	0 – 63	1 or -1 or 0	67	-1
	64 – inf	1 or -1 or 0		1
RRs	0 – 90	1 or -1 or 0	85	1
	91 – inf	1 or -1 or 0		-1
HR	0 – 45	1 or -1 or 0	48	-1
	46 – 100	1 or -1 or 0		1
	101 – 150	1 or -1 or 0		-1
	151 – inf	1 or -1 or 0		-1

After data transformation, the discrimination coefficient is computed (according to above mentioned formula).

The network consist of 5 layers i.e. That from each learning pattern it will be choose 5 feature with maximal value of discrimination coefficient. It is shown in following table:

Table 4. Chosen 6 features with the best value of discrimination for output classes

Feature no.	Patterns with output classes (A,B,C,Z)																			
	Z	Z	Z	C	Z	Z	Z	B	Z	Z	Z	Z	A	Z	Z	A	Z	Z	Z	B
1	0,22	0,12	0,12	0,26	0,12	0,12	0,22	0,26	0,12	0,12	0,12	0,22	0,04	0,12	0,12	0,46	0,12	0,12	0,22	0,26
2	0,22	0,12	0,12	0,26	0,12	0,12	0,22	0,26	0,12	0,12	0,12	0,22	0,04	0,12	0,12	0,46	0,12	0,12	0,22	0,26
3	0,13	0,37	0,37	0,59	0,37	0,37	0,13	0,21	0,37	0,13	0,37	0,37	0,74	0,37	0,37	0,74	0,13	0,37	0,37	0,29
4	0,13	0,37	0,37	0,59	0,37	0,37	0,13	0,21	0,37	0,13	0,37	0,37	0,74	0,37	0,37	0,74	0,13	0,37	0,37	0,29
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0,04	0	0	0	0,04	0	0	0	0	0,04	0,13	0,13	0,04	0	0	0	0,04

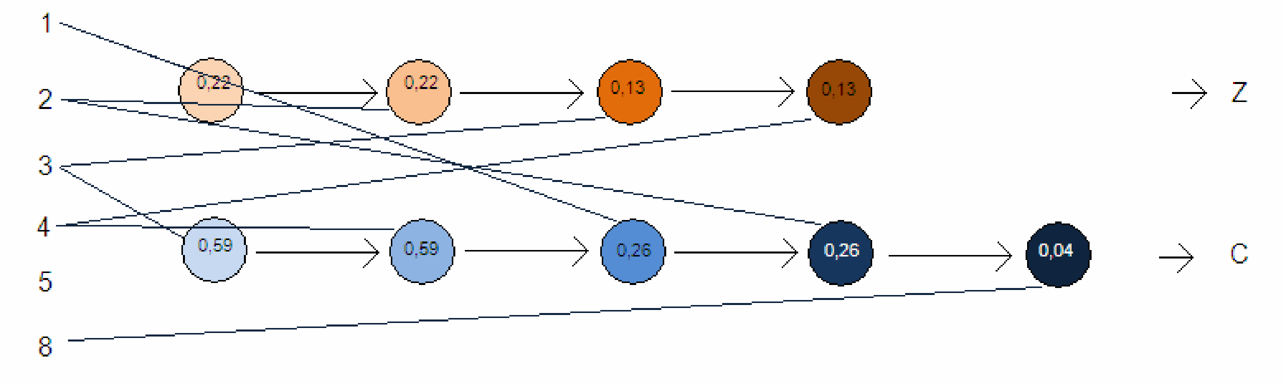

Fig.1. Example of building neural network topology based on marked value of discrimination from Table 4.

Assuming that the network has 5 layers, we obtain the following network topology:

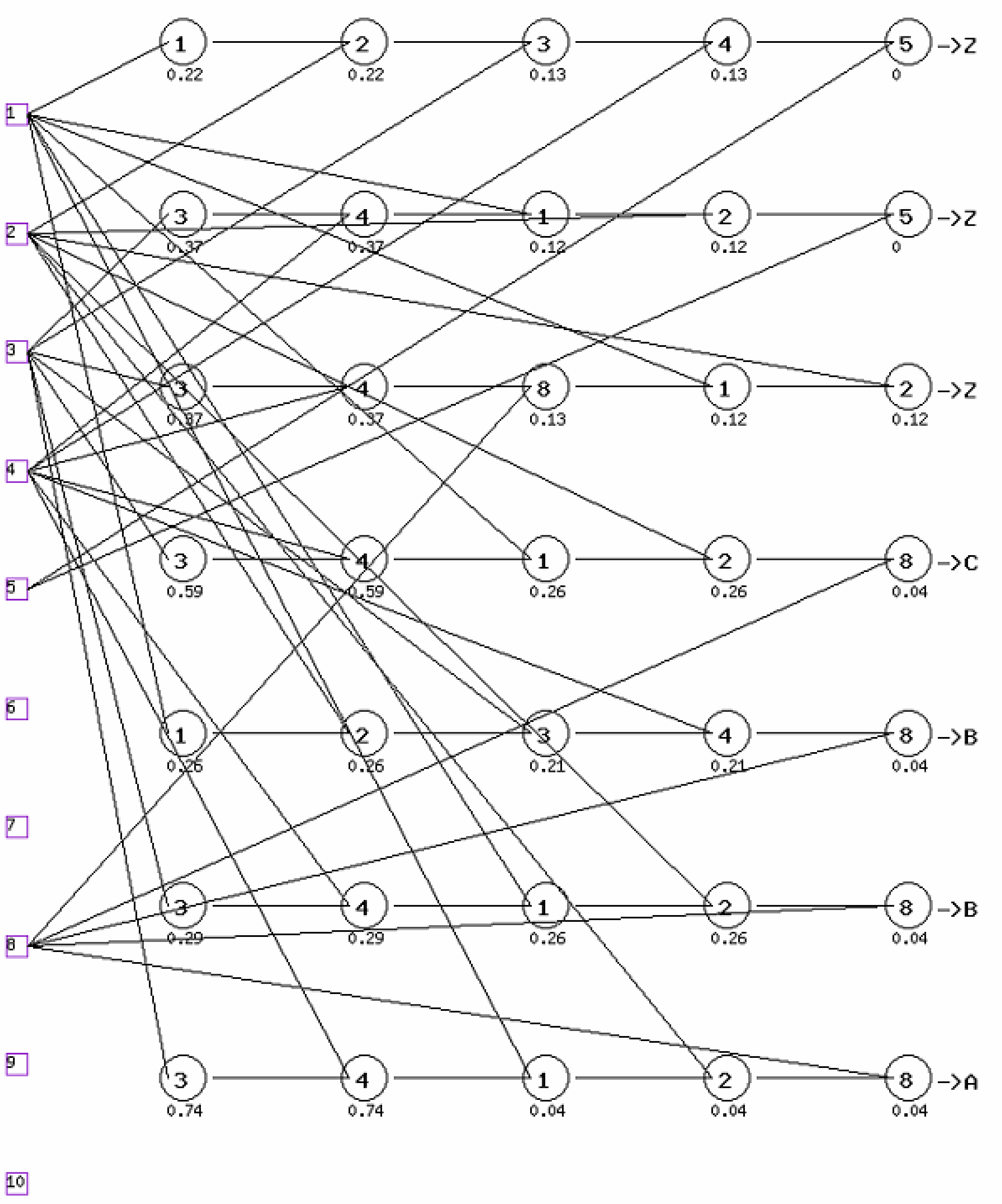

Fig.2. Ready to work neural network.

On this figure it can be seen, that some features very well discriminate particular patterns, e.g. feature 1,2,3,4,5,8. These features described sex, age, RRs (from 0-90) and HR (46-100). At the same time other features are not considered at all, e.g. 6,7,9,10. It is easily to seen that during construction of network, features with maximal value of discrimination coefficient have priority over other feature. Each following layer consists of neurons with decreasing value of coefficient.

Choice of the best discriminative features is strictly connected with set of patterns and mathematical formula for coefficient of discrimination. For other data set or different formula for discrimination coefficient, it can be different choice of the best discriminative features. But always it will choice of such features, which univocally classified particular patterns.

7. REFERENCE

Choice of the best discriminative features is strictly connected with set of patterns and mathematical formula for coefficient of discrimination. For other data set or different formula for discrimination coefficient, it can be totally different choice of the best discriminative features. But always it will choice of such features, which univocally classified particular patterns. Choice of the best discriminative features is strictly connected with set of patterns and mathematical formula for coefficient of discrimination. For other data set or different formula for discrimination coefficient, it can be totally different choice of the best discriminative features. But always it will choice of such features, which univocally classified particular patterns.

BIBLIOGRAPHY

[1] S. WOJCZYK, A. POPA, M. LIZIS, Idea Samooptymalizujacych Sieci Neuronowych SONN budowanych na bazie metrycznej przestrzeni cech wejsciowych, rozdział 29 w monografii: Bazy danych- nowe technologie, Wydawnictwo Komunikacji i Łacznosci, Warszawa 2007, 315-324.

[2] S. WOJCZYK, A. POPA, M. LIZIS, Samooptymalizujace siatkowe sieci neuronowe, rozdział 30 w monografii: Bazy danych - Nowe technologie, Wydawnictwo Komunikacji i Łacznosci, Warszawa 2007, 325- 333.

[3] A. HORZYK, R. TADEUSIEWICZ, Self-Optimizing Neural Networks, Advances in Neural Networks - ISNN 2004, Yin F., Wang J., Guo C. (eds), (Springer-Verlag, Berlin Heidelberg, 2004) 150-155.