ACTA UNIVERSITATIS AGRICULTURAE ET SILVICULTURAE MENDELIANAE BRUNENSIS

Volume LIX 21 Number 2, 2011

EVOLUTION OF INSURANCE COMPANY SERVICE QUALITY SURVEY, USING SELF-LEARNING NEURAL NETWORK

V. Konečný, O. Trenz, D. Dvořáková

Received: December 17, 2010

Abstract

KONEČNÝ, V., TRENZ, O., DVOŘÁKOVÁ, D.: *Evolution of insurance company service quality survey, using self-learning neural network*. Acta univ. agric. et silvic. Mendel. Brun., 2011, LIX, No. 2, pp. 149–154

The objective of the paper is to demonstrate the abilities and possible approaches to classification of set of objects using self-organizing maps. As the objects, clients of an insurance company that made an agreement regarding mandatory insurance of motor vehicles were selected. The opinions of the clients and their overall satisfaction reflected in responses to presented answers.
The clients were classified into three groups. The first two contained satisfied clients (i.e. good clients for the company), the last group contained clients that could potentially switch to the competitors. Subsequent analysis enabled discovering the reasons of low customer satisfaction and critical factors of losing the least satisfied clients.
For the analysis of the responses (one hundred fifty-one) and the insurance company, experimental model of self-organizing map realized at the Department of informatics was used. Used experimental model has proved very effective software tool.

insurance company, neural network, self-learning, classification, class representative, plane projection

The increase in a number of users with powerful personal computers allows using of unconventional methods for decision-making support. This trend is evident both in collecting necessary data and their processing that is complexly reflected in applications of business intelligence (BI) using also methods and tools of artificial intelligence (AI).

Artificial neural networks are often used tools by the AI. Self-learning neural network, Kohonen's network (Kohonen, 2001) respectively, providing also a graphical display of position of *good*, *worse*, and *critical* customers has been chosen for analysis of customer satisfaction survey of an insurance company from several options.

1 METHODS AND RESOURCES

1.1 Input information

Information about the views of clients of insurance companies has been obtained from responses to the issues raised in the questionnaires posted on the website. The selection of respondents was carried out randomly even if the target group was clients of only one particular insurance company.

Questions and answers (except the names of insurance companies) selected for the views of clients together with a numerical rating necessary for creation of a client position vector in the input space of neural network, are as follows.

1. *Insurance company client*: the names will not be published.
2. *The reason for choosing insurance company*: trust in the insurance company (1), quality of service (0), low insurance cost (3), recommended by friends (3), comfortable mediation (4), recommended by advisor (3) vehicle purchase (5), other response (3).
3. *Duration of insurance*: ten years or more (0), five to 10 years (2), two years to 5 years (3), less than two years (5).
4. *Satisfaction:* Very satisfied (0), satisfied (1), other reviews (3), dissatisfied (4) very dissatisfied (5).

5. *What clients prefer:* fast processing (0), the limit of insurance (3), the size of policy (5), quality of assistance services (1) communication (2) nothing (3), other (5).
6. *What can be improved:* nothing (0) unanswered (3), other (3), the size of policy (3), the quality of assistance services (4), communication (4), the limit of insurance (5).
7. *Satisfaction with the expertise of insurance workers:* very satisfied (0), satisfied (1) nothing (2), dissatisfied (4).
8. *Claims*: none (0), one (2) two or more (4).
9. *The course of settling the claim:* no event (0), fast (0), comfortable (1), other (3), lengthy (4), complicated (5).
10. *Information about news:* telephone (0), letter or e-mail (2), during a visit to an insurance company (1), unknown (3), the insurance company does not inform (5).
11. *Provided benefits*: discounts on a car insurance (0) vouchers (1), bonus transfer (2), a gift (3), unknown (3) I do not know (3), none (5).
12. *Transfer to another insurance company*: loyalty to one (0), when buying a car (3), unknown (3), I watch the offers each year (4) when I was dissatisfied (5).
13. *Use of accident insurance*: yes (0), the first years (2), unknown (3), windshield (3), not (5).
14. *Information about CIB (*Czech Insurers' Bureau*):* from the insurer (0), from other sources (3), unknown (3) news (5), the client has already paid the fine (5).

Responses were obtained from a total of one hundred fifty-one respondents from nine different insurance companies. Most respondents were clients of the five insurance companies. Clients of remaining insurance companies were classified within one collective group. The following Table I lists the evaluated responses of a target insurance company.

The listed evaluation of clients' response reflects the view of insurance personnel. All evaluated data form an input vector of the respondent in the structure: *x(Respondent identifier, a list of evaluated responses).*

Generated respondent identifier consists of his/her serial number within respondents of one insurance company and an insurance company number.

1.2 Information Processing

In general, the point is to use an appropriate method for classification of objects (Konečný, Trenz, 2009b), (Kohonen, 2001). In the field of an artificial intelligence, we cannot forget either the multi-layer

I: *Evaluated responses of respondents*

Client	Question number												
	2	3	4	5	6	7	8	9	10	11	12	13	14
x300	**3**	**5**	**1**	**3**	**5**	**1**	**0**	**0**	**2**	**0**	**3**	**3**	**5**
x301	3	2	1	2	3	2	0	0	2	0	3	0	3
x302	1	0	1	2	3	0	0	0	2	0	0	5	3
x303	3	5	1	3	5	1	0	0	5	0	3	3	0
x304	1	0	1	5	3	0	2	1	2	0	0	5	3
x305	4	2	1	0	4	1	2	1	2	0	5	2	0
x306	**3**	**2**	**1**	**5**	**5**	**1**	**0**	**0**	**1**	**3**	**5**	**3**	**3**
x307	**5**	**0**	**4**	**3**	**3**	**4**	**4**	**4**	**5**	**5**	**3**	**3**	**3**
x308	**5**	**2**	**4**	**5**	**3**	**1**	**0**	**0**	**5**	**5**	**3**	**0**	**5**
x309	0	3	1	0	4	0	2	0	2	2	0	3	3
x310	1	3	1	0	5	1	0	0	2	2	0	0	3
x311	**0**	**5**	**1**	**3**	**3**	**0**	**4**	**1**	**2**	**0**	**3**	**5**	**3**
x312	0	0	1	2	3	1	2	1	1	0	0	0	0
x313	0	3	0	0	3	0	2	0	1	0	0	0	3
x314	**3**	**0**	**1**	**0**	**4**	**1**	**4**	**0**	**5**	**5**	**5**	**3**	**5**
x315	1	0	1	1	3	0	4	0	2	0	0	0	3
x316	**0**	**5**	**1**	**0**	**4**	**1**	**2**	**4**	**5**	**0**	**3**	**3**	**3**
…	…	…	…	…	…	…	…	…	…	…	…	…	…
x344	**0**	**3**	**1**	**0**	**3**	**1**	**0**	**0**	**2**	**0**	**5**	**5**	**5**
x345	5	3	1	0	3	1	2	1	2	0	3	2	0
x346	0	0	0	0	3	0	4	1	5	0	0	0	3
x347	3	5	1	3	3	1	0	0	2	0	0	5	3
x348	3	5	1	3	3	1	0	0	0	0	0	3	0
x349	5	2	1	0	3	0	2	1	1	0	0	0	5
x350	3	2	1	1	3	1	4	1	2	0	0	0	3

neural network which learns itself from the models, i.e. a teacher, or we can implement a fuzzy classification system. For a given problem, however, it seems easiest to use self-learning neural network.

It is a double layer neural network. The first layer consists of input neurons which merely repeat the input values. Each input layer neuron is connected to each output layer neuron with weights $w_1, w_2, \dots w_n$ representing output neurons' vectors W.

This network is working so that every combination of inputs forming the vector $X = (x_1, x_2, \dots, x_n)$ can activate only the output neuron W^i for which

$$|X - W^i| < |X - W^k|, \forall k. \tag{1}$$

This means that if X and W are the position vectors of points of n-dimensional vector space, then the only output activated is the one which represents the point closest to the point represented by the vector X. The activated neuron is then called the winning neuron and its vector W is being corrected in the learning process so that it gets even closer to the input vector X. The correction can be formalized in a relationship

$$W^* = W + \alpha(X - W), \tag{2}$$

where W^* is a new vector of the winning neuron and is the learning coefficient (correction) whose value decreases with increasing number of stages of learning. Within the used neural network model,

$$\alpha(E) = \frac{\alpha_0}{1 + \left(\dfrac{E}{E_{max} K_\alpha \sqrt{3}}\right)^2}, \tag{3}$$

$\alpha_0 < 1$ is the initial learning coefficient, E_{max} – the maximum number of stages of learning, $K_\alpha \in <0.2;0.6>$ – the coefficient indicating the position of the inflexion point of a function $\alpha(E)$ with respect to the E_{max}.

To comply with the principle of distance correction, it is necessary to carry out a correction of neurons adjacent to the winning one, as well. Correction is made the same way as for winning neurons but with a coefficient of learning $\beta(E,d)$, instead. Moreover, this coefficient drops continually with distance d from the winning neuron.

$$\beta(E, d) = \begin{cases} \alpha(E) & , d_0 \geq d > 0, \\ \alpha(E)\exp\left(-\left[\dfrac{d - d_0}{T\beta}\right]^2\right) & , d > d_0, \end{cases} \tag{4}$$

where d_0 is the radius of the neighbourhood with a coefficient of learning $\beta(E, d) = \alpha(E)$. The parameter d_0 together with parameter

$$T_\beta = (d_{ext} - d_0)\sqrt{2} = K_\beta R_0 - d_0)\sqrt{2} \tag{5}$$

enable better modelling of the course of the function $\beta(E, d)$. Distance d_{ext} determines the position of an inflection point of this function. Using K_β, this point is entered due to an actual radius of neighbourhood R_0.

The neighbourhood of the winning neuron is defined in the used model by relation

$$R_0(E) = R_m \times \exp\left(-\frac{E}{T_{R_0}}\right), T_{R_0} = \frac{E_1}{\ln(R_m)}, \tag{6}$$

where $E_1 = K_{R_0} \times E_{max}$.

R_m is an initial radius of the neighbourhood and E_1 indicates a number of learning stages after which $R_0<1$. K_{R_0} allows entering E_1 due to E_{max}.

The process of vectors' correction of output neurons is a basis for a process of self-learning neural network. Input vectors are entered in a random order and the winning vectors in the stage of learning are getting closer to the input vectors.

After a sufficient number of steps, the output vectors W will be equal with input vectors X or vectors representing centre of gravity of points with a unit weight represented by two or more inputs. Evidently close vectors can have one common output.

If necessary, output neurons are usually arranged in one- or two-dimensional formation. In this case, self-learning neural network realizes objects' display with position vectors X from n -dimensional input space to output space represented by neurons with a principle of maintaining distance. This means that closer objects within an input area will be also close in an output space. Detail of the issue of self-learning is mentioned in the monograph (Konečný, Trenz, 2009a).

If the neural network would be used only to project objects of n -dimensional space into two-dimensional, the result will provide only their relative positions. In order to view classification classes, it is necessary to provide representatives, i.e. objects for which the distance of class representatives is smaller than representatives of other classes. Expert may designate representatives, either by selecting one object or group of objects of each class or it is possible to use a network of one-dimensional arrangement of output neurons for determination representatives of a self-learning neural network. In the case of the model $(1 \times N_R)$, N_R is a number of representatives.

If an expert decides to define a representative set of k objects for each class, it is possible to determine the i[th] coordinate of a class representative as a centre of gravity of this group of objects by the relationship

$$x_i^k = \frac{1}{k} \Sigma x_i^k. \tag{7}$$

RESULTS

Classification of respondents according to affiliation to insurance companies 10, 30, 40, 50, 60, and 70, in which insurance companies with a small number of respondents are combined, does not cause any serious trouble. Insurance company number distinguishes adequately among the customers and, therefore, six representatives of separate classes can be determined not only by relation (7), but also by self-learning of neural network with output into a map *(1 x 6)*, see Fig. 1. In both cases, the coordinates of representatives should be at worst very close. For example,

(10.00, 2.30, 3.50, 0.80, 2.70, 3.30, 1.00, 0.80, 0.00, 2.30, 0.80, 2.90, 1.90, 1.70)

are calculated coordinates of a representative *r10* for an insurance company 10, and

(10.0003, 2.3036, 3.5063, 0.8009, 2.7098, 3.3983, 1.0002, 0.7963, 0.0000, 2.2950, 0.7970, 2.9009, 1.9007, 1.7055)

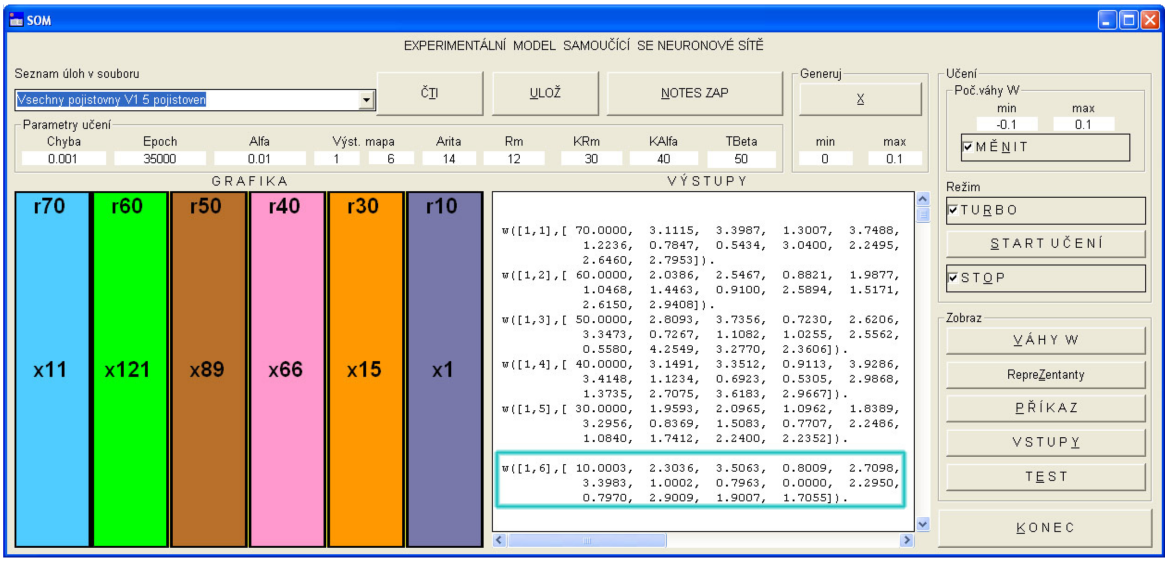
are the coordinates determined by a model.

An appearance of the graphical user interface of an experimental model is shown in Fig. 1 Vectors of output neurons, respectively the representatives of six sets of clients, are in a text box. Colour frame accentuates vector *w(1,6) = r10* which represents an insurance company 10.

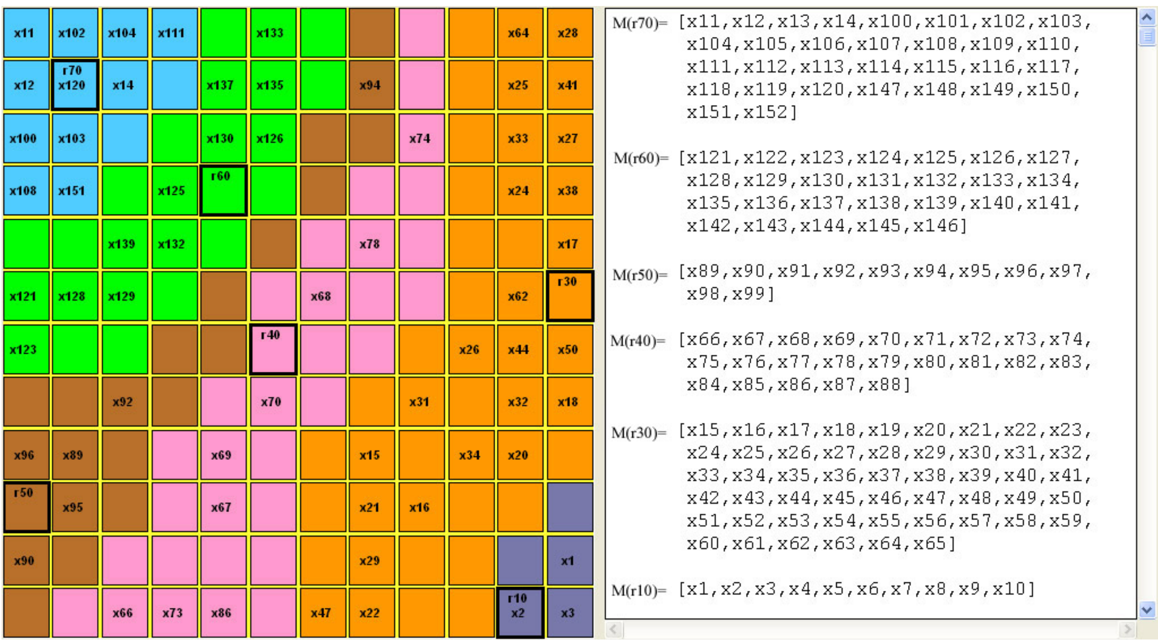
Client classification according to insurance companies processed by a model of the self-learning neural network is presented in a graphic form, a map format *(12 x 12)*, and as a list of clients, in Fig. 2. Typical clients are depicted close to a representative.

A separate file was created with a skipped coordinate *insurance company number* for final exploration of clients of the insurance company 30. A set of respondents of this insurance company contains elements with identifiers *x300, x301,…, x350*.

Client evaluation of an insurance company will be based on a comprehensive influence of responses



1: *Outputs of representatives of insurance companies*



2: *Display of all respondents according to insurance companies* to thirteen questions. A classification will be implemented into three sets which represent *good (satisfied)*, *worse (or less satisfied)*, and *critical* clients of an insurance company.

Self-learning with an output *(1 x 3)* can provide coordinates of vectors of three representatives for required sets and the respondent distribution to these sets according to rule of a minimum (Euclid's) distance from a set representative. Methods of determining classes of representatives are further discussed in the monograph (Konečný, Trenz, 2009).

The described method of distribution (Fig. 3, set specifications M(r1), M(r2) and M(r3)) does not yet determine a meaning of particular sets. This can be accomplished by an expert analysis of representatives' attributes. In this case, it will be easier to find a set of satisfied clients according to position of display of an ideal client who pertain a zero evaluation of responses to survey questions. The ideal client was used in the neural network model as a test object, and therefore he is surrounded by a circle in Fig. 3. Importance of other sets is determined due to distances of their representatives from a representative from a set of *satisfied* clients.

According to calculations made by software of the experimental model, it was found that representative *"r1"* is the closest one to the ideal client *"ide"*, representative *"r2"* and *"r3"* follow in the listed order (see Fig. 3). Specifically: distance (ide, r1) = 5.4718, distance (ide,r2) = 7.3665, distance (ide,r3) = 9.7518. This means that the *"r1"* represents *satisfied* clients, *"r2" – less satisfied* clients, and *„r3" – critical* clients. Lists of elements of these sets are shown in Fig. 2, outputs (1 x 3). Mutual placement of objects within the sets and their colour resolution is noticeable on graphic part of the Fig. 3.

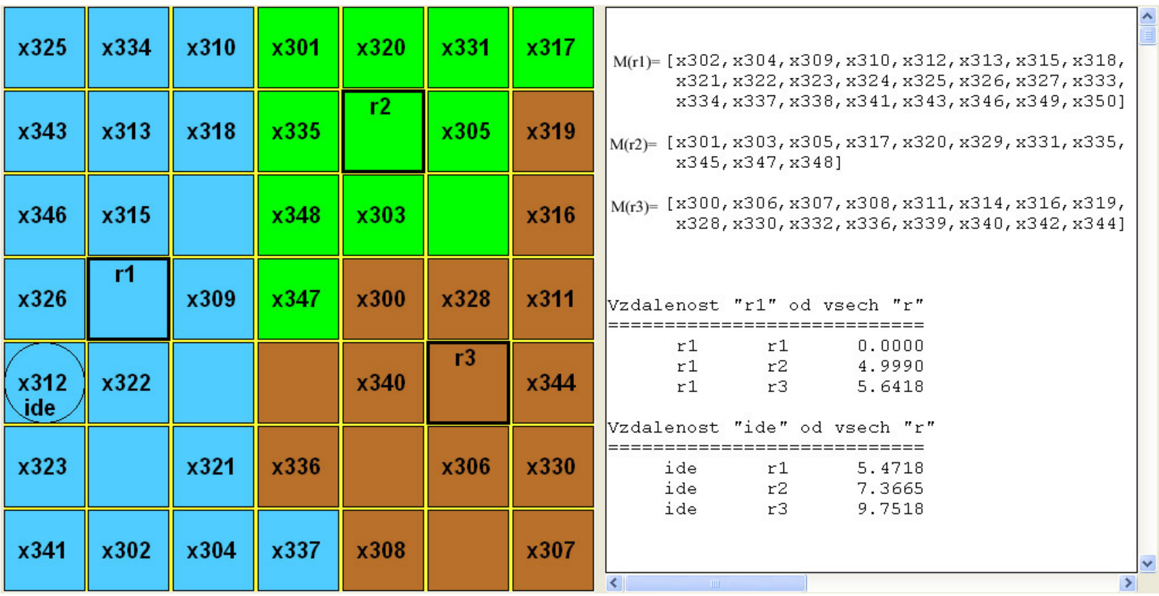
DISCUSSION

Sets of *r*1 and *r*2 representatives do not present a threat. The set marked *r*3 of 16 clients represent the greatest dissatisfaction clients, which is 31.7% of the total number of 51 respondents.

Critical set of clients is highlighted in Table I by colour. It is easy to make certain that this classification is mostly caused by a negative response from few questions. What these questions are can be identified by the coordinates of sets' representatives in Table II.

Position of each set's representative (centre of gravity) is determined by its coordinates. In the case of set *r*3, the centre of gravity is strongly influenced by negative responses to question numbers 6, 10, 12, 13, and 14 (highlighted). They are the answers to the following questions: What can be improved (No. 6), information on the news (No. 10), transfer to another insurance company (No. 12), use of accident insurance (No. 13), information about CIB (No. 14).

When asked *"what can be improved"* most clients of *r*3 set demands reduction of cost and limit of insurance but often also occurs as dissatisfaction with the quality of services.



3: *Classification into three groups*

II: *Coordinates of representatives of target insurance company*

	2	3	4	5	6	7	8	9	10	11	12	13	14
r1	1.2372	0.7354	0.9677	1.3505	3.1274	0.6246	1.8547	0.7105	1.8409	0.7052	0.0000	1.6592	2.0867
r2	2.5379	3.8155	0.8160	1.8162	3.2702	0.7293	0.5412	0.2706	1.8186	0.1807	2.9165	2.3575	0.5532
r3	2.6185	2.8703	1.5038	2.6269	**3.5568**	1.2571	1.6229	1.1729	**3.2662**	2.3948	**3.6258**	**3.0582**	**3.7010**

III: *Representatives' coordinates of clients of all insurance companies*

	2	3	4	5	6	7	8	9	10	11	12	13	14
r1	1.5448	1.5770	0.7599	1.0078	3.3474	0.7366	1.6344	0.6949	1.7750	0.7907	0.3451	1.6748	1.8093
r2	2.6843	3.8459	0.9481	3.5054	3.4011	0.9471	0.6423	0.5391	2.8152	0.3894	3.3494	3.0529	2.7128
r3	3.6277	2.8533	1.6309	3.7388	3.1955	1.5805	1.4258	1.0048	3.7113	4.6576	2.8781	3.4670	3.5627

Improving *information on the news* at the current potential of CIB should not be a bigger problem. *Transfer to another insurance company* will depend on an approach to dealing with the consequences related to answers for the 2nd question. The answers for *use of accident insurance* do not provide detailed information but we can assume that the cause will be its cost. Improving awareness about CIB is more or less linked with information on news so that ensuring correction should not be problematic. Specific steps depend on a management of an objective insurance company but we can assume that an increase in care of existing customers could significantly reduce proportion of "critical customers".

Table III lists the coordinates of representatives of all respondents disregarding an insurance company. *r3* is, similarly to the previous case, a representative of critical customers. Deflection in centre of gravity of a representative for this group from satisfied clients is influenced, as well as for an independently researched insurance company, by answers to questions: *information on the news (No. 10), use of accident insurance (No. 13)*, and *information about CIB (No. 14)*.

The reason for choosing insurance company (2) belongs to other factors of worse evaluation of this group by clients. Adverse reactions are caused likely by clients of small insurance companies with shorter duration so that neither *confidence* in the insurance company nor the *quality of services* can be evaluated. Other variants of the answer (*comfortable mediation, recommended by advisor, purchase of a car*) have higher "penalization" and thus shift a respondent to the group of worse or critical customers. An independently surveyed insurance company 30 could use this knowledge and present quality of services and insurance companies selling vehicles to the future owners in more appropriate way.

Answers *the limit of insurance* and *the size of policy* prevail in the questions *what clients prefer* (5) which is probably the main reason why they selected an insurance company where they established a compulsory Road Traffic Act Insurance.

Most clients correspond negatively to the questions of *provided benefits* (11) as the average response value (see Table III) is 4.6576 and the answer *no* has penalization 5. This knowledge is useful for gaining new clients as well.

As it is evident from the procedure, the used neural network allows easy and fast evaluation of client survey of insurance companies. A self-learning neural network with a number of outputs equal to a number of required sets should suffice for the determination of representatives and full specification of elements of required sets. However, the projection of objects (clients) of a multidimensional space to a plane with a colour designation of the sets gives a better idea of their position in relation to representatives of classes.

SUMMARY

The paper presents the use of self-learning neural network on a real example. The used application enables easy identification of representatives of classification classes of clients of several insurance companies and more detailed analysis of clients of the selected insurance company. The weight of each attribute in the classification of customers is relatively easy to be determined according to representatives of classes. The knowledge gained can be used to identify your ideal client, as well as critical insurance clients. Such knowledge can be used to create insurance offerings.

REFERENCES

KONEČNÝ, V., TRENZ, O., 2009a: Rozhodování s podporou umělé inteligence. Folia Mendelovy zemědělské a lesnické univerzity v Brně, Folia II, 2009, 8, ISSN 1803-2109, ISBN 978-80-7375-344-3.

KONEČNÝ, V., TRENZ, O., 2009b: Classification of Companies with Assistance of Self-Learning Neural Network. Agricultural Economics. ISSN 0139-570X.

KOHONEN, T., 2001: Self-Organizing Maps. 3rd edition, Springer Verlag Berlin. ISBN 3-540-67921-9.

Address

doc. Ing. Vladimír Konečný, CSc., Ing. Oldřich Trenz, Ph.D., Dana Dvořáková, Ústav informatiky, Mendelova univerzita v Brně, Zemědělská 1, 613 00 Brno, Česká republika, e-mail: konecny@mendelu.cz, xtrenz@node.mendelu.cz, xdvora16@node.mendelu.cz