DOI: https://doi.org/10.59503/29538009-2023.si-1-33

USAGE OF ARTIFICIAL NEURAL NETWORKS IN MARKETING RESEARCH

Gayane Ghukasyan*
Yerevan State University
Ph.D., Associate Professor
gayaneghukasyan@ysu.am

Petros Petoyan
IMR LLC Founder and CEO
petros@imr.am

Gayane Bakhshyan
IMR LLC CDO
gayane@imr.am

Abstract: Any organization, in order to carry out successful activities in the conditions of strong competition and uncertainty of the environment, must be constantly informed of its customers, familiar with their constantly changing requirements, understand what is its target audience and strive to maximize their satisfaction level. For this purpose, the implementation of marketing research is very important, because it allows to collect multi-faceted information about customers, their requirements, the attitude of customers towards the given organization and the product or service issued by the latter, and other important factors that directly or indirectly affect the competitiveness of the organization and on profitability [1].

Modern economics puts forward new methods, the use of which in the implementation of marketing research allows the process to be carried out more efficiently, quickly and accurately. Artificial neural networks (ANNs) are among those methods. They allow you to identify extremely complex dependencies, perform analyzes for large amounts of data, solve complex formalization problems, etc. The article presents the analysis of the results of the survey conducted for the purpose of assessing the satisfaction of consumers with the level of water supply using artificial neural networks and offers recommendations aimed at solving the above problems.

Keywords: marketing research, artificial neural networks, machine learning, customer segmentation, object classification, multilayer perceptron

JEL code: M31

Research goals: the purpose of the research is the application of artificial neural networks in marketing research for the purpose of processing and analyzing the results of surveys (water supply sector) and based on this, presenting recommendations on the effectiveness of the use of ANNs in marketing research.

Research novelty: multilayer neural network (Multilayer Perceptron) was built for the purpose of processing and analyzing the results of marketing research inquiries.

Introduction

Due to the high degree of competition in the modern concept of marketing, more and more companies start to pay attention to the analysis of the needs of the target audience. In order to face the

growing competition, any organization must study the ever-changing needs and preferences of customers, understand which is its target audience and strive to maximize the level of customer satisfaction. Therefore, companies strive not only to make their products of higher quality, but also to adapt them to the needs and interests of a specific consumer. Here, the quality and accuracy of the information collected about consumers is of utmost importance [2]. In practice, traditional marketing research is accompanied by mathematical methods, the advantages of which include the ability to analyze the situation and make optimal decisions. Mathematical methods imply the construction of a formalized description of the object under study and its use for further analysis.

Currently, modeling through using artificial neural networks is gaining wide popularity for research [3], [4]. Neural networks are used in various fields (medicine, marketing, banking, real estate, etc.) in order to solve many problems of classification, clustering, prediction, control. In particular, [6] is dedicated to the modeling of creditworthiness assessment of commercial bank customers using ANN, and [7] to the clustering of RA banks using Kohonen neural networks.

Artificial neural networks allow to identify extremely complex dependencies between variables, to perform analyzes for large-scale data, to solve complex formalization problems, etc. Networks can be used for qualitative data presented in quantitative form, and therefore also in marketing research for the processing of survey results.

By analyzing current customers, the neural network enables them to be segmented and identify which customers are beneficial

to the company and which factors influence their behavior the most. The implementation of such analyzes is of great importance, because in order to increase the competitiveness and efficiency of their activities, enterprises must clearly understand who their consumer is and what internal and external factors are most important for him.

Due to the rapid change of information, it is becoming more and more difficult for companies to track consumer behavior, which leads to an increase in market research costs. The integration of neural networks in research activities will allow not only to save the costs of consumer research, but also to free up time for specialists, as well as to obtain the most reliable model that will enable effective communication with customers and maintain the company's competitiveness.

Research results

In the work, artificial neural networks were used in marketing research for the purpose of processing and analyzing survey results (using the example of the water supply sector, based on anonymous survey data of 1206 respondents).

For the assessment of water supply quality, ANN models were built, which belong to the class of teacher-taught problems.

In order to evaluate the satisfaction of consumers with water supply services, various factors are important, which have been grouped into three groups: assessment of customer satisfaction with the services provided by water supply companies, assessment of the process of responding to consumer requests, assessment of the change in the water supply situation compared to the previous

year. Artificial neural networks were built to evaluate the satisfaction of consumers in each request and to solve the problems that arose. For the sake of brevity, the article provides an assessment of the change in the water supply situation compared to the previous year.

Consumers' satisfaction with water supply (dependent variable, the answers are: has the situation improved, remained the same compared to last year, or has it worsened) depends on various factors (independent variables): water purity, safety, usefulness, taste of water, pressure of water flowing from the tap, water supply schedule.

At first, in order to get an idea of the relationship between the dependent and independent variables, cross tabulations were built in advance.

In order to evaluate the influence of all independent variables on the dependent variable, a multi-layer perceptron was constructed using the Neural Network tool of the SPSS package. The input variables of the network are tap water purity, water safety, taste, pressure, and water supply schedule. The constructed network has 15 input neurons (one of three possible responses can be chosen for each of the independent variables), excluding the constant, 1 hidden layer with 8 intermediate neurons and 3 output neurons for each group of customers. The activation function of hidden layer neurons is the Sigmoid function, and the activation function of the output neurons is the Identity function. The error is estimated by the Sum of Squares function (the sum of the squares of the differences between the target and estimated values).

For the purpose of neural network training, non-intersecting subsets were separated: training, test and independent. In the constructed model, the training set includes 75.2% of all data, the test set 12%, and the independent set 12.8%. The network is trained with data from the training set. The test sample is intended to check the accuracy of the model and estimate the error, and the independent set is used to test the performance of the already constructed network on completely new, unfamiliar data.

The neural network is represented by Figure 1.

Estimates of connection weights are given in the Parameter Estimates table (Figure 2). In the Hidden Layer 1-Input Layer part of the table, the weights from the inputs to the neurons of the hidden layer are given, and in the Output Layer-Hidden Layer 1 part, the weights from the hidden layer to the neurons of the output layer are given. The system selects these weights automatically before training. If such a choice of weights leads to the fact that the value of the error function is greater than a certain threshold value, then with the help of the error Backpropagation algorithm, the weights are adjusted in order to reduce the error.

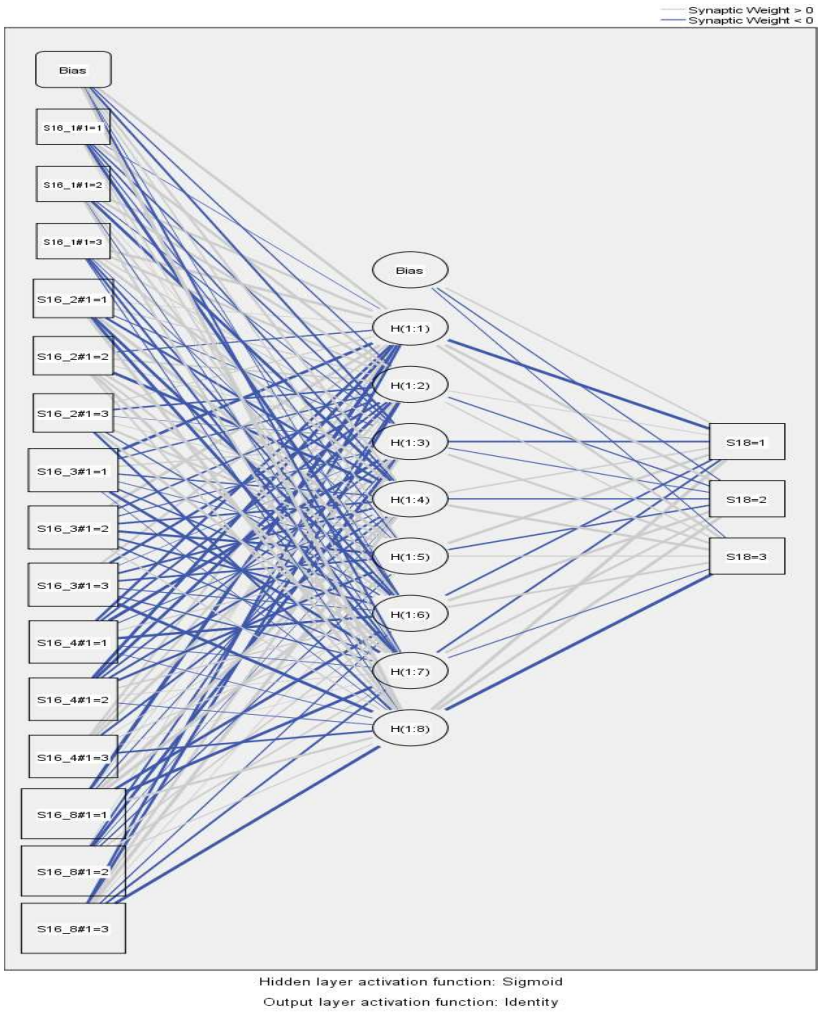

Figure 1. A Neural Network Built for Marketing Research[1]

[1] Composed by authors

Parameter Estimates

Predictor		Predicted									
		Hidden Layer 1							Output Layer		
		H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)	H(1:6)	H(1:7)	[S25=1]	[S25=2]	[S25=3]
Input Layer	(Bias)	-.110	.113	-.099	.016	-.159	-.225	.015			
	[S20=1]	-.377	-.600	-.142	.074	-.123	.743	-.370			
	[S20=2]	-.199	-.126	.806	-.057	-.297	-.481	-.062			
	[S20=3]	-.415	.240	.067	-.093	-.471	-.320	-.334			
	[S20=4]	.491	.396	.319	.063	.023	.248	-.473			
	[S20=5]	.383	-.791	.018	.019	-.107	-.191	-.225			
	[S20=6]	.174	-.262	-.140	.501	-.307	-.508	-.444			
	[S20=7]	.149	.198	.224	-.034	.501	-.038	.264			
	[S20=8]	.013	.418	.377	-.051	.277	.159	.071			
	[S20=9]	.222	.274	.869	.443	-.244	-.582	.021			
	[S23=1]	-.064	-.662	-.481	-.537	-.414	-.101	-.987			
	[S23=2]	.105	-.399	.501	.294	-.411	.006	-.267			
	[S23=3]	.293	.094	-.268	-.075	-.152	.493	-.396			
	[S23=4]	.403	.206	.426	.452	-.055	.492	.821			
	[S23=5]	.202	.438	-.016	-.101	.224	-.769	.700			
	[S23=666666]	.345	-.086	.074	.353	-.062	-.252	-.406			
Hidden Layer 1	(Bias)								.290	1.022	-.469
	H(1:1)								-.395	-.380	-.054
	H(1:2)								-.690	.225	.759
	H(1:3)								-.364	.150	.869
	H(1:4)								-.378	-.437	.278
	H(1:5)								-.123	-.101	.184
	H(1:6)								.444	-1.012	.261
	H(1:7)								-.718	-.150	1.020

Figure 2. Estimates of link weights in the network[2]

Table 1 shows the results of correct and incorrect classification for each category of the dependent variable (%). In the trained sample, the neural network correctly classified 59.4% of the "situation improved" responses, 88.5% of the "remained the same" responses, and 19.7% of the "deteriorated" responses. Overall accuracy in the trained sample is 75.4%.

In the test sample, 64.9% of "the situation has improved" answers, 87.6% of "remained the same" answers, and 40% of

[2] Composed by authors

"deteriorated" answers were correctly classified. The prediction accuracy in the test sample is 77.9%. In the independent sample, 65.2% of the "situation has improved" responses, 89.2% of "remained the same" responses, and 28.6% of "deteriorated" responses were correctly classified. Overall, the prediction accuracy in the independent sample is 78.8%.

Table 1. ANN classification results
Source: Our research results

Sample	Observed	Predicted
Training	improved	59.4%
	remaind the same	88.5%
	deteriorated	19.7%
	Overall Percent	75.4%
Testing	improved	64.9%
	remaind the same	87.6%
	deteriorated	40.0%
	Overall Percent	77.9%
Holdout	improved	65.2%
	remaind the same	89.2%
	deteriorated	28.6%
	Overall Percent	78.8%

The model evaluated the importance of each of the independent variables in the process of predicting the results. It is presented in the form of an Independent Variable Importance diagram (Figure 3). In the ranking process in the model, the highest importance was given by the consumer to the water supply schedule (100.0%),

followed by the importance of water taste (75.1%), tap water pressure (72.7%), safety (58.1%) and cleanliness (47.2%).

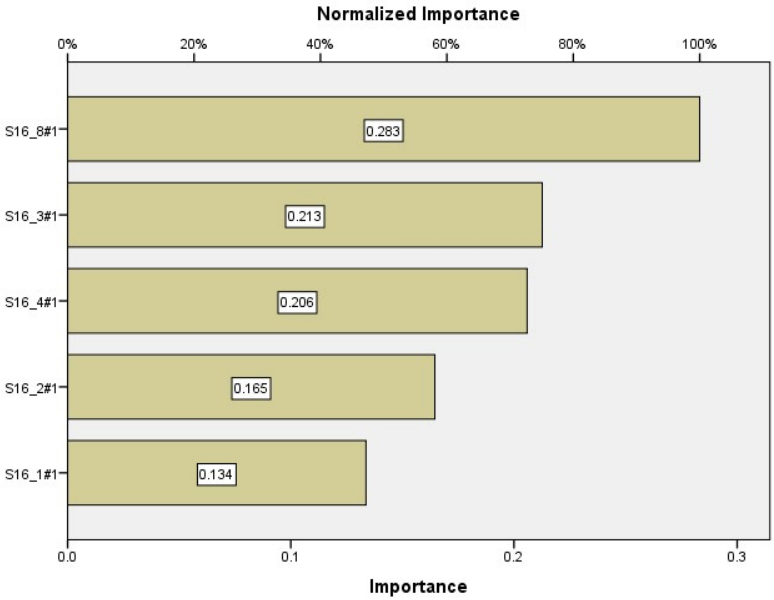

Figure 3. Independent Variable Importance diagram[3]

Multinomial regression model: Multinomial regression model was constructed to conduct a comparative analysis. As is known, multiple logistic regression is a general case of the logistic regression model where the dependent variable has more than two categories, each of which can be measured on a Nominal or Ordinal scale, and the independent variables can be Ordinal or Scale [5]. The model is designed to solve classification problems, it estimates

[3] Composed by authors

the probability of the dependent variable falling into a certain class and is used to classify new objects.

We have a set of possible g values of the dependent variable $\{1, 2, \ldots, g\}$, where the probability of accepting the k-th value is:

$$P[y = k] = p_k, k = 1, 2, \ldots, g.$$

The probabilities are given by the following formula:

$$p_1 = P[y = 1] = \frac{1}{1 + \sum_{h=2}^{g} exp(\alpha_h + x\beta_h)}$$

$$p_k = P[y = k] = \frac{exp(\alpha_k + x\beta_k)}{1 + \sum_{h=2}^{g} exp(\alpha_h + x\beta_h)}, k = 2, \ldots, g,$$

where the coefficients α_k and β_k are estimated by the Maximum Likelihood Method [5].

As in the case of the model built with a neural network, the results of the classification are presented here with the Classification table (Table 2). All variables in the model are represented in three categories.

For the "improved" category, 53.1% of the data were correctly classified, for the "remained the same" category - 91.9%, for the "worsened" category - 21.8%. In general, the model correctly classified 76.1% of the data, which is close to the result obtained with the use of neural networks.

Table 2. Classification results by Multinomial Regression model

Source: Our research results

Observed	Predicted
improved	53.1%
remaind the same	91.9%
deteriorated	21.8%
Overall Percent	76.1%

Now let's consider the second problem: the satisfaction of customers with the services provided by water supply companies. For this purpose, we will use the assessment of the change in the water supply situation by the customers depending on the following variables:

- ✓ in case of questions, means of contacting the water supply company or its representative,
- ✓ the speed of solving problems by the customer support center,
- ✓ courtesy of the representative/s of the water supply company
- ✓ provision of information about interruptions by the water supply company,
- ✓ the level of providing general information by the water supply company.

The constructed network has 15 input neurons, 1 hidden layer with 8 intermediate neurons and 3 output neurons for each group of customers. The activation function of hidden layer neurons is the Sigmoid function, and the activation function of the output neurons is the Identity function.

The overall accuracy in the trained sample is 72.7%, in the test sample - 64% and in the independent sample - 78.4%.

The model presents the importance of each factor in the neural network prediction process. The most important is the courtesy of the representatives of the water supply company (100%), then the speed of solving problems (84.4%), the level of receiving information about interruptions (76.7%), the level of providing general information (60.3%), and the company means of contacting representatives (59.8%).

Conclusion

Modern economics puts forward new methods, the use of which in the implementation of marketing research allows the process to be carried out more efficiently, quickly and accurately. Among those methods is the methodology of artificial neural networks.

Consumers' satisfaction with the water supply situation was assessed in the work. A multilayer neural network was built to process and analyze survey results. The classification results were compared with the results obtained by the Multinomial Regression model. The factors that have the greatest impact on consumer satisfaction have been identified. The obtained results confirm the effectiveness of using artificial neural networks in the process of developing and analyzing the results of marketing research. In addition to the fact that neural networks are able to find complex dependencies, solve problems that have quite a lot of variables, they also allow identifying the factors that have the most significant

impact on the studied variable and focus the attention of specialists on those factors.

As a result, it is revealed for which consumer segment and what kind of changes should be implemented in order for the organization to gain a competitive position.

References:

1. ***Kotler, Ph.,*** Fundamentals of Marketing: Short course, Moscow, 2007, p. 646.
2. ***Tyurin, D. V.,*** Marketing research: Organization and implementation in the company, Moscow, 2016, p. 341.
3. ***Haykin, S.,*** Neural networks: a complete course, 2nd edition. Per. from English. M. ed. "Williams", 2006, p. 1104.
4. ***Hagan, M. N., Demuth, H. B., Beale, M. H.,*** O. De Jesus. Neural Network Design, 2014, p. 800.
5. ***Magnus Y. R., Katishev P. K., Peresetskiy A. A.*** Econometrics. Basic course: the tutorial. Sixth edition. Moscow: Delo.2004, p. 576.
6. ***Ghukasyan, G., Ghazaryan A.,*** Clustering and classification of RA banks using the Kohonen neural network, "Development of economy and society. "Challenges and opportunities of the 21st century" conference material, 2017, pp. 409-416.
7. ***Ghukasyan, G., Khalatyan, A.,*** Modeling creditworthiness assessment of commercial bank customers using artificial neural networks, Finance and accounting, 2019(2), pp. 70-84.

ԱՐՀԵՍՏԱԿԱՆ ՆԵՅՐՈՆԱՅԻՆ ՑԱՆՑԵՐԻ ԿԻՐԱՌՈՒՄԸ ՄԱՐՔԵԹԻՆԳԱՅԻՆ ՀԵՏԱԶՈՏՈՒԹՅՈՒՆՆԵՐՈՒՄ

Գայանե Ղուկասյան
Երևանի պետական համալսարան, ֆ.-մ.գ.թ., դոցենտ

Պետրոս Պետոյան
«Այ Էմ Առ» ՍՊԸ հիմնադիր տնօրեն

Գայանե Բաշխյան
«Այ Էմ Առ» ՍՊԸ զարգացման տնօրեն

Բանալի բառեր - մարքեթինգային հետազոտություն, մեքենայական ուսուցում, արհեստական նեյրոնային ցանցեր, հաճախորդների սեգմենտավորում, օբյեկտների դասակարգում, բազմաշերտ պերսեպտրոն

Հոդվածում ներկայացվում է արհեստական նեյրոնային ցանցերի կիրառմամբ ջրամատակարարման մակարդակից սպառողների գոհունակության գնահատման նպատակով անցկացված հարցման արդյունքների վերլուծությունը:

Ջրամատակարարման ծառայությունների որակի գնահատման նպատակով կառուցվել են ԱՆՑ մոդելներ, որոնք պատկանում են ուսուցչով ուսուցանվող խնդիրների դասին:

Ջրամատակարարման ծառայությունների նկատմամբ սպառողների վերաբերմունքի գնահատման համար դիտարկված գործոնները խմբավորվել են երեք խմբում. ընկերության մատուցած ծառայությունների որակի գնահատում, սպառողների խնդիրներին արձագանքելու գործընթացի գնահատում, նախորդ տարվա նկատմամբ ջրամատակարարման

իրավիճակի փոփոխության գնահատում: Բացահայտվել են այն գործոնները, որոնք առավել էական ազդեցություն են գործում սպառողների գոհունակության վրա: Արդյունքում աշխատանքի արդյունավետությունը բարձրացնելու նպատակով ընկերությունը պետք է իր ուշադրությունը և գործունեությունը կենտրոնացնի այդ գործոնների վրա:

Ջրօգտագործման ոլորտի օրինակով կատարված հետազոտությունը փաստում է արհեստական նեյրոնային ցանցերի կիրառման արդյունավետությունը մարքեթինգային հետազոտություններում: Արհեստական նեյրոնային ցանցերի ինտեգրումը մարքեթինգային հետազոտությունում հնարավորություն է ընձեռում հաճախորդների պահանջմունքների, նախասիրությունների վերաբերյալ հավաքագրված բազմակողմանի տեղեկատվության վերլուծության գործընթացն արդյունավետ, արագ և ճշգրիտ կերպով իրականացնել: Այն թույլ է տալիս նաև նվազեցնել հետազոտության ծախսերը և ժամանակը:

Վերլուծելով ընթացիկ սպառողներին՝ նեյրոնային ցանցը հնարավորություն է տալիս վերջիններիս սեգմենտավորել և բացահայտել, թե ո՞ր հաճախորդների հետ համագործակցությունն է հանդիսանում շահավետ ընկերության համար, և ո՞ր գործոններն են առավել մեծ չափով ազդում նրանց վարքագծի վրա:

Submitted: 25.10.2023; Revised: 20.11.2023; Accepted: 22.11.2023