Nonlinear Prediction Models in Data Analysis

Željko Račić[1], Zoran Ž. Avramović[2], Đuro Mikić[2]

[1]University of Banjaluka, [2]PanEuropean University APEIRON Banjaluka

zeljko_r@hotmail.com, zoran.z.avramovic@apeiron-edu.eu, djuro.d.mikic@apeiron-edu.eu

A General Survay

https://doi.org/10.7251/JIT2002106R UDC: 537.533.35:510.5

Abstract: The modern entrepreneurial sensibility of the company's business implies directing the right information to the appropriate parts of the company at the right time. That is why it is necessary to digitalize processes as much as possible and make the organization "intelligent", and its human resources, to the greatest extent, the knowledge workers. The application of neural networks, i.e. nonlinear prediction models, enables systematic analysis of data in the function of evaluating the behavior of the system. Neural networks are a powerful tool, especially for forecasting trends and forecasting based on historical data. The grouping method, i.e., the k-mean value algorithm, is used as a precursor to neural networks.

Keywords: neural networks, Back-propagation neural network, grouping methods, k-mean algorithm.

Introduction

In the last fifty years, technology has changed the world. Science has advanced more than in the entire history of mankind, while the total human knowledge, compared to the previous period, has increased several times, and the time required to increase knowledge is getting shorter. Information technology and its use are the most frequently mentioned topics in the business world. The reason for the great interest is its accelerated development and opening of great opportunities for business applications. Information and communication technology is changing the ways people work and live, and is changing the organization and way of doing business in modern companies. Those who do not adapt to these changes - either individuals or companies, will bring into question their existence and successful functioning in the newly created business and technological environment. Knowing some models and methods can fill a glass of prejudice, it can keep us in one place without allowing us to look at the problem from another angle. John Maynard Keynes defined it in 1936 with the saying: The difficulty lies not so much in developing new ideas as in escaping from old ones [5]!

Today we are witnessing that, even in the time of the pandemic, we must adapt the company's business only to electronic forms of business. Those companies that have been trained for this type of business, work smoothly and fulfill the set goals. That is why it is necessary to know the possibilities provided by modern information technologies, and the context and business environment in which they operate in today's companies. This primarily refers to the ever-present gap between technology and business-oriented people, who so often have completely different visions of what information technology is for a company and how to make full use of its possibilities. According to scientific research, the average company uses less than 10% of the collected data. When we add external data (competition, macroeconomic indicators, etc.), we see that the situation is far from ideal. The amount of data we produce is growing exponentially. The fact is that today we are "suffocating" in the data, and we demand more and more knowledge. Obviously, there is a great need in the business world to analyze this data. Methods for analyzing this data exist, and some of them will be described in this paper. There are a number of so-called main and generally accepted methods, but there are also a number of methods from other areas that cannot be categorized. Some methods are constructed with the help of elements of linear algebra, information theory, mathematics and other

fields. This paper will describe the methods that are expected to contribute to the goals of research work, and these are also the most important representatives from the family of a certain group of methods, such as neural networks and grouping methods. In economic research, we first prepare the data by classifying them into groups and then analyze them using one of the methods listed in the paper.

Grouping Methods

Grouping can be defined as finding similar entities (with common features) within data. Groups are formed by the process of sharing a set of data, where group affiliation is defined on the basis of similar characteristics (for example, gender, age, district). The grouping algorithm aims to find similarities within the given data set, using the selected attribute set. Depending on the specific method, groups can be defined in the following ways:

- exclusive: each element belongs exclusively to one of the groups;
- overapping: an element can belong to several groups at the same time;
- probabilistic: the element belongs to each of the groups with a certain probability;
- hierarchically structured: with a rough division of elements at the highest level, which can then be more finely structured at lower levels.

The aim of this analysis is to create a procedure for grouping elements into groups (clusters), so that each group is, as much as possible, homogeneous with respect to the grouping variables. The first step in grouping is to choose a measure of similarity, followed by deciding on the type of grouping technique we will use (hierarchical or non-hierarchical). The third step is to select the type of grouping method for the selected technique (e.g., the centroid method in a hierarchical grouping technique). In the fourth step, we make a decision based on the number of groups. Finally, an interpretation of the results follows. Hierarchical methods do not require a priori knowledge of the number of groups. Grouping can be performed using multiple methods: centroid method, single bond method, full connectivity meth-

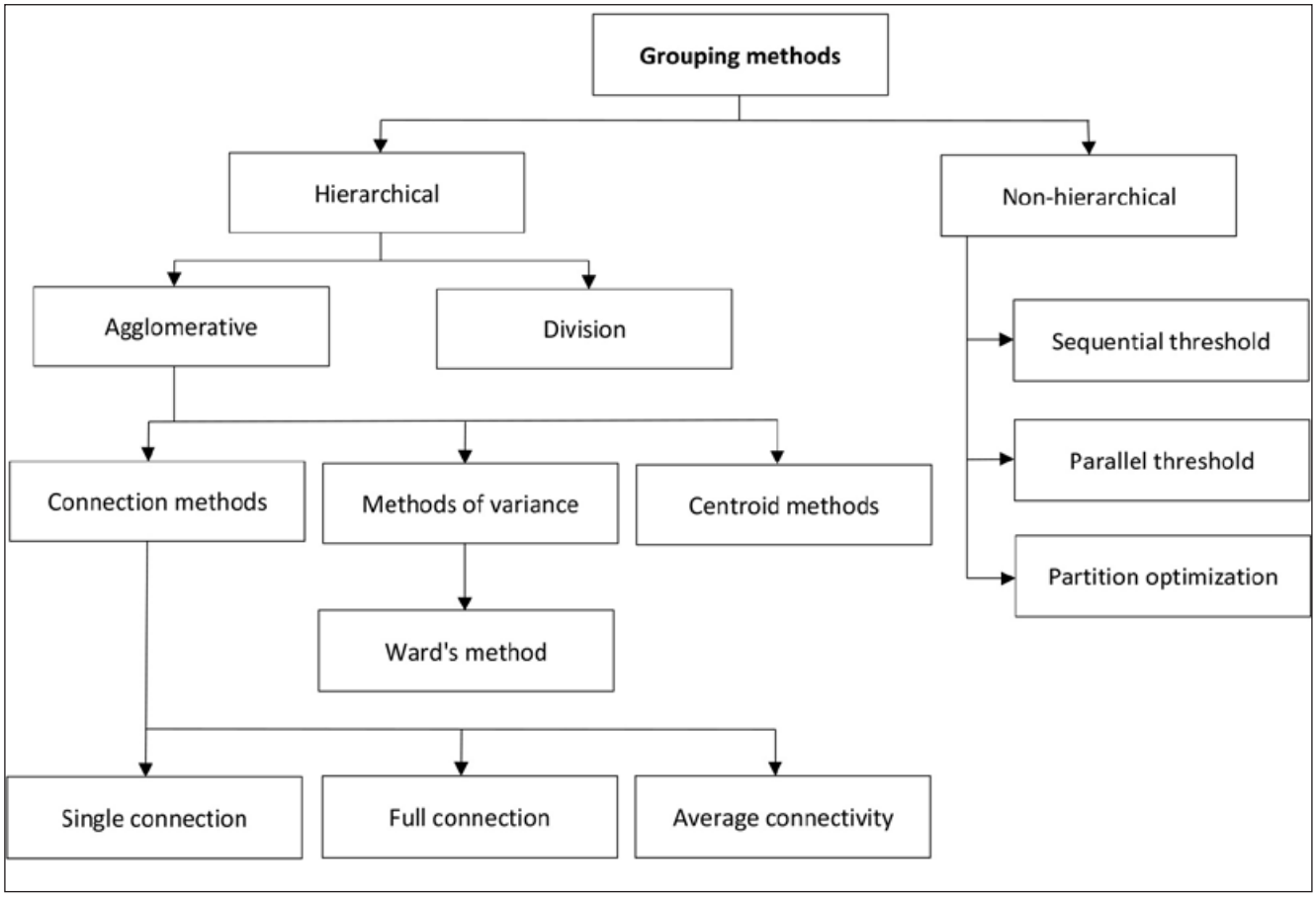

Figure 1: *Grouping method types*

od, average connectivity method, Ward method. The disadvantage of hierarchical methods is that we cannot redistribute an assigned element to another group. Therefore, the hierarchical method is used for the purpose of detecting groups, and the resulting solution is further refined by non-hierarchical methods.

K-mean algorithm

There are a number of grouping algorithms, but the most well-known grouping algorithm is the k-mean algorithm, which, using distance estimation functions and centroids, creates groups in an iterative procedure.

The main characteristic of this method is the division of the basic population into "k" groups, where each of the groups contains "n" similar elements. The similarity of the elements is estimated by the algorithm based on the distance function. Initially, the number of desired groups is set and centroids are defined for each group. In the iterative procedure, each group is joined, using the distance function, by the most similar elements from the population (the calculation is done based on the calculation of the distance of each element of the population from the central values ("centroids"). After each cycle (iterations), the new mean values of each group (centroids) are recalculated, and a new iterative cycle is entered, until the centroid values become stable. This process is illustrated by the following scheme:

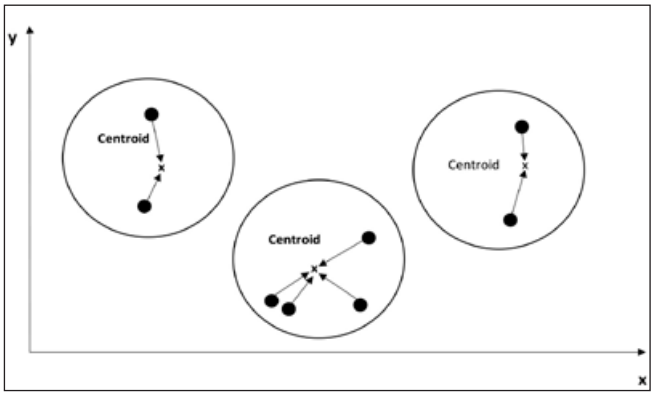

Figure 2: *Simplified scheme of the k-means algorithm [8]*

The following characteristics need to be considered when preprocessing grouping data:

1. perform standardization on the series of attribute values that participate in the analysis process (thus neutralizing the influence of variables with the highest range of values on the analysis results); for example, we perform grouping using the k-means algorithm values where the analysis process processes two variables: the customer's age, can range from 12-100, and the amount of purchase for six months in a retail facility, can range from 0-4000; in case the variables are not standardized, the variable purchase amount would have a significant impact on the analysis due to the large interval (range) of values in relation to the customer's age;
2. transform the values of non-numerical variables into numerical ones (it is necessary to design a numerical system that will best interpret descriptive values taking into account the meaning of grouping algorithms); this means that, for example, when numbering municipalities (cities), their territorial distance will be taken into account - for example, Prijedor will be assigned a numerical value that is closer to Banja Luka than is the case with, for example, Trebinje: in this way the algorithm will recognize the similarity at the territorial level between the districts.

The main disadvantage of the grouping method stems from the fact that the decision maker determines the number of groups before conducting the analysis, while in practice it is necessary to conduct several iterative grouping processes to form the optimal number of groups.

This method is rarely used as a target method, but rather as a method for "field testing" and introducing analysts to the data (for example, initial market segmentation). Due to the impossibility of interpreting the results of the analysis on the formed groups (for example, a group of buyers of healthy food, a group of buyers of luxury products), other methods of data analysis (most often, neural networks) are applied to them.

Neural Networks

The main idea of neural networks comes from neuropsychology. The knowledge about the behavior

of the nerve cells of the human brain, which function on the principle of activation, i.e. increasing the potential at the synapses, is used. The functioning of the human brain is not fully known, but according to existing knowledge, the following is considered vali [7]:

1. information processing in the human brain is realized through a network of millions of simple process units, which are called neurons (the human brain contains about 10^{11} neurons);
2. each neuron is a simple processor: it receives signals from a large number of other neurons, combines them and sends signals to other neurons;
3. knowledge is distributed in the human brain through a large number of connections between neurons.

Existing biological neural networks are incomparably more complex than the mathematical model used in practice. In the mathematical model of the neural network, the basic unit is designed on the model of a biological neuron. Units combine inputs into a single result (usually a summation function), which is then redirected to a transformation function that calculates the output value and usually takes the values 0 and 1. The combined and transfer function form the activation function of neurons. Figure 1 shows the mentioned relations.

Neural networks differ from each other in terms of the learning approach they use. Learning approaches can be divided into:

1. supervised learning: the neural network is given the right answers together with the input quantities, i.e. real output quantities, where the system itself determines the weighting coefficients (for example, when a student learns a new lesson with the help of a teacher [7]);
2. supervised learning: the system receives only input values, without knowing the output values, and is expected to detect or recognize some patterns of behavior in the data itself (for example, a child's learning that is exposed to repeated situations; a typical example is the Kohonen neural network (self-organizing maps([7].

Although there are hundreds of neural network architectures, here we will only talk about the most popular architecture: multilayer perceptron or Back-propagation neural network (BNP network). The BNP network is organized into three or more layers of neurons (input, hidden, and output), where data travels in one direction. Each neuron is actually an independent logical unit that receives information in the form of stimuli at its receptors that have different levels of significance (depending on the weighting factor), and accordingly the stimuli

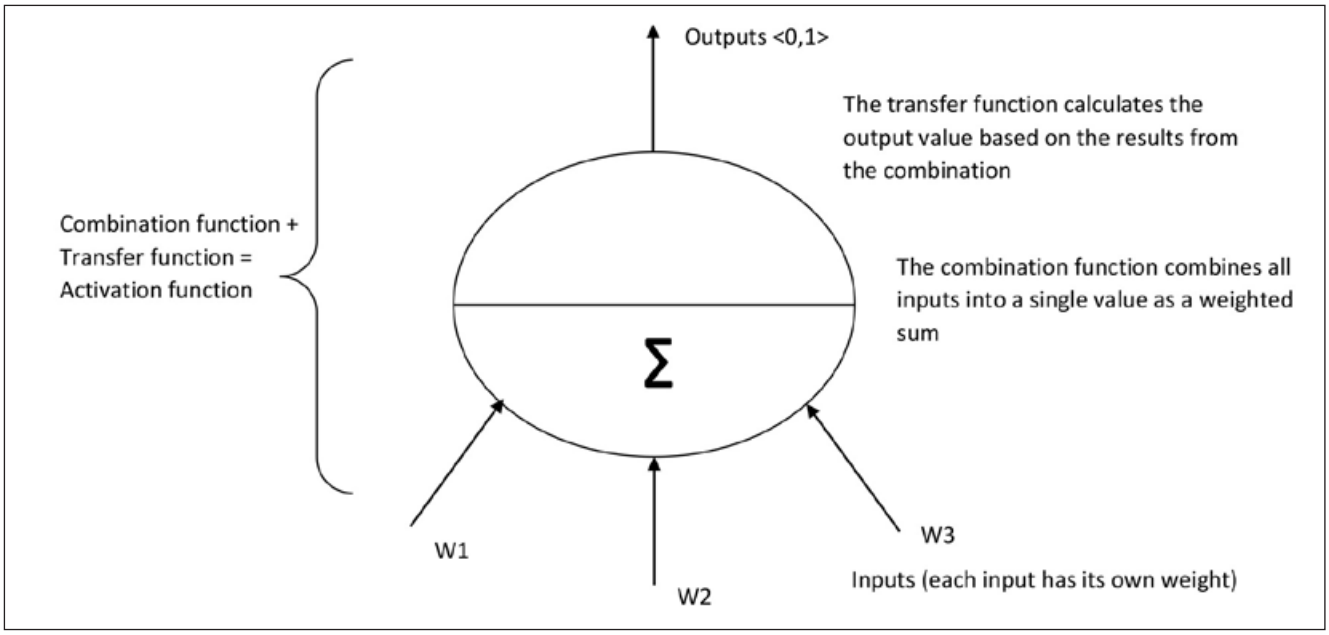

Figure 3: *Mathematical model of neurons [6]*

become stronger or weaker. If the sum of all stimuli is greater than the sensitivity threshold of the neuron, then the neuron sends (transmits) a pulse at the output that serves as an input pulse to the neurons in the next layer. It is important to note that there is no single neural network model that would apply to all types of problems. The characteristic of each (so far) developed model is that it has certain advantages when applied in a specific area. The advantage of neural networks is that they can be easily applied and implemented in a large number of parallel processors, with each processor simultaneously performing its own calculation.

The application of neural networks, ie nonlinear prediction models, enables modeling of large and complex problems. Neural networks are a powerful tool, especially for forecasting trends and predicting based on historical data. They serve to answer questions such as, for example: If the price of product X decreases by a certain percentage, by how much will the demand for that product increase? The problem with neural network-based models is the determination of appropriate values for weighting coefficients and sensitivity thresholds. Existing algorithms solve this problem by looking for a local minimum in nonlinear space. Another problem with neural networks is that the knowledge in them is not descriptive, that is, it cannot be translated into a human-readable form, so they should be viewed as black boxes. This is also the reason why they are not acceptable in cases where the description of the acquired knowledge is of great importance, and when analysts actually want to get an answer to the question of how the network came to the results.

Conclusion

In their work, modern companies are increasingly oriented towards the integration of their business activities, and in general a more comprehensive and extensive view of their business processes. However, without the support of modern software applications, as well as information and communication technology, such an approach is not possible.

Special attention should be paid to the construction of the model and the choice of the appropriate technique. In the business field, it is important that the obtained models are understandable to those who use the results of the analysis to make business decisions. The importance of interactive and visualization techniques, which have the possibility of visualization, is especially emphasized here, as very often a picture speaks more than words. The use of complex, highly parameterized models may not always be the best solution. The use of analysis techniques that can have their own visual presentation shortens the total time of the analysis, as well as undertaking activities based on the results obtained in the analysis.

References

[1] Amor, D. The E-Business Revolution, Prentice Hall, New York, 2002.
[2] Han, J., Kamber, M. Data Mining: Concepts and Techniques, Morgan Kaufmann Publishers, 2000.
[3] Han, J., Altman, R, Kumar, V, Mannila, H, Pregibon, D, 2002. Emerging Scientific Applications in Data Mining, Communications of ACM 45, 2002.
[4] Haris, D. B. Creating a Knowledge Centric Information Technology Environment, Knowledge Media, New York, 2003.
[5] Keynes, J.M.: The General Theory of Employment Interest and Money, MacMillan&Co Ltd, London, 1964.
[6] Linoff, G., Berry, J.A.: Mastering Data Mining-The Art and Science of Customer Relationship Management, John Wiley & Sons, Inc., 2000.
[7] Soldić-Aleksić, J. 2001. Inteligentni sistemi za poslovno odlučivanje, Ekonomski fakultet, Beograd.
[8] Panian, Ž., Klepac, Ž. Poslovna inteligencija, Masmedia, Zagreb, 2003.
[9] Račić, Ž., Stanić, S. Ekonomsko-matematički modeli i metode, Univerzitet u Banjoj Luci, Ekonomski fakultet, 2019.

Submitted: September 23, 2020
Accepted: November 21, 2020.

About the authors

Željko Račić is an Professor at the Department for Quantitative Analysis and Informatics, Faculty of Economics and Faculty of Electrical Engineering, University of Banja Luka, Bosnia and Herzegovina. He has gained his education at the Faculty of Economics in Banja Luka where he has obtained his Bachelor and PhD degree, and at the Faculty of Economics in Belgrade where he has received his Master's degree. He is an author of numerous books and scientific papers and a member of several international organizations.

Zoran Avramović is the rector of Pan-European University APEIRON, Banjaluka, Republic Serpska, Bosnia and Herzegovina. Born in Serbia (Yugoslavia) on September 10, 1953. He graduated from the School of Electrical Engineering of the University of Belgrade. At this School he received a degree of a master, and then a PhD of technical sciences University of Belgrade, 1988. He was awarded the Gold Medal of the Academy of Electrotechnical Sciences of the Russian Federation (AES RF) for his achievements in the field of electrical engineering (2005). So far, he has published 365 scientific and professional papers.

Đuro Mikić defended his doctoral dissertation entitled "Statistical modeling of business success parameters" in 2003 at the Faculty of Economics, University of Banja Luka. Since 2013, he is a full professor at the Faculty of Business Economics of the Pan-European University Apeiron Banja Luka, in the scientific field of operational research and statistical analysis. During his academic and scientific career of almost 40 years, he is the author of 39 scientific and professional papers published in eminent domestic and international scientific journals, and 10 books in the field of statistical analysis, decision theory, management, and systems theory.

For citation

Željko Račić, Zoran Avramović, Đuro Mikić, Nonlinear Prediction Models in Data Analysis, *JITA – Journal of Information Technology and Applications Banja Luka*, PanEuropien University APEIRON, Banja Luka, Republika Srpska, Bosna i Hercegovina, JITA 10(2020) 2:106-111, (UDC: 537.533.35:510.5), (DOI: 10.7251/JIT2002106R), Volume 10, Number 2, Banja Luka, December 2020 (69-128), ISSN 2232-9625 (print), ISSN 2233-0194 (online), UDC 004