

SCIENTIFIC REVIEW

DEFINING OF NECESSARY NUMBER OF EMPLOYEES IN AIRLINE BY USING ARTIFICIAL INTELLIGENCE TOOLS

Dragan Petrovic²⁸, Mirjana Puharic²⁹, Edita Kastratovic³⁰

ABSTRACT

In modern business, uncertainty and risks are increasing, and the available time is not enough to make the right decisions. The consequence of such a dynamic environment is the creation of flexible organizations and efficient managers who are ready to quickly respond to market demands using modern technologies. In this paper the model for preliminary estimation of number of employees in airline by using of artificial intelligence tools. It is assumed that the tools of artificial intelligence can be applied even for complex tasks such as defining the number of employees in the airline. The results obtained can be used for planning the number of employees, ie. planning the necessary financial investments in human resources, and may also be useful for a preliminary analysis of the airlines that choose to do restructuring or plan to increase/decrease the number of operations. Results were compared with those obtained by regression analysis.

KEY WORDS: Artificial intelligence tools, aircraft, airline company, human resources, management and planning, analysis

JEL: M12

UDC: 331.108.22:656.7

656.7

COBISS.SR-ID 272230412

²⁸ Directorate of civil aviation, Banja Luka, Republika Srpska dragan.petrovic@bhdca.gov.ba

²⁹ Faculty of Information Technology and Engineering, University Union-Nikola Tesla, Belgrade, Serbia

³⁰ Faculty of Business Economics and Entrepreneurship, Belgrade, Serbia

INTRODUCTION

Planning is a continuous process that reflects changes in the environment and is adapted to them. Business management in each decision determines the economic performance of the organization in the first place. Their existence and continued engagement in the company management confirms the achieved economic results. (Kastratović, 2008). Management and planning of human resources is one of the key problems faced by any company, regardless of whether it is production, providing of services or both. Defining the required number of employees is a complex issue that depends on many factors, so it is necessary to have a good knowledge of industry in which company operates, the entire business process, technology, and techniques of business, then the infrastructure that includes technical resources, facilities and operational business organization. In accordance with the requirements of stakeholders that influence the changes in the environment, and therefore the organization, the human resource management policy is changing. In doing so, emphasis is placed on the necessary qualifications, attitudes and expectations of employees. (Kastratović, 2008).

Data mining is the process of finding, relationship, rules and regularities in the data set. Data mining could be used for creation of predictive models, finding significant factors in the data sets, definition of the database of rules for the expert system. Data Mining represents automated analytic process designed to effectively and efficiently explore a large database with the aim of discovering, validation and extraction of valuable, hidden information concerning until then unknown facts of cases and models. Some of the basic techniques and methods of Data Mining of databases are neural networks, fuzzy logic, statistical methods, genetic algorithms and genetic programming.

As the rapid growth of quantity of data in many areas of human activity goes beyond the power of man for processing and analyze such a large number of data by conventional means, there was a need for techniques of automatical analysis of data, and on this new generation of techniques and tools for analysis was focused study in the field of Knowledge Discovery in large databases.

In (Petrović, 2000), (Petrović, 2001), (Petrović, 2001), and (Puharić et al, 2013) the application of genetic programming and neural network (abbreviated NN) was shown in various fields of aviation from the design of the aircraft, through the management of the financial indicators. This paper is a logical extension of these studies.

It should be noted that the data mining that uses artificial intelligence tools is widely used not only in the civil environment, but can have military applications in almost all aspects. The results of this study can also be applied to the air force, but due to the lack of reliable data, the research was applied to civil aviation. The results can be particularly useful in the part relating to military transport aircrafts.

ARTIFICIAL NEURAL NETWORKS

There are two categories of NN: artificial and biological NN (Peterson et al, 1992) and (Kriese, 2007). A representative of biological NN is the neural system of living beings. Artificial NN have the structure, function and information processing similar to biological NN, but they are artificial creations. NN in computer science is a highly connected network of elements that handle and process data. One of the important characteristics of NN is their ability to learn on a limited set of examples. In this paper, when we talk about NN, we refer primarily to the "artificial NN" (abbreviated ANN), because it is mostly about the model of NN, implemented on computers. NN represent a system composed of a large number of simple processing elements. Such systems are capable of collecting, storing and use of experimental knowledge. There is no single definition of NN. However, NN could be defined as follows: A NN network is a system composed of many simple processors (units,

neurons), each of which has a local memory which memorizes the data that are being processed. These units are connected through communication channels (connections). Unit handles only its local data and inputs they receive via the connections. Limitations of local operators can be removed during training. NN are collections of mathematical models that simulate some of the observed properties of biological neural systems and withdraw similarities with adaptive biological learning. They are made up of a large number of interconnected neurons (processing elements) which, like the biological neurons, linked with their connections that contain a permeable (weighting) coefficients, which are similar to the role of the synapses. Learning in biological systems is performed by regulating synaptic connections that connect the axons and dendrites of neurons. Learning by examples of typical events is achieved through training over accurate input-output data sets by repetition of algorithm and adjusting bandwidth (weight) ratios of connections (synapses). Most NN has some kind of rules for "training", by which the coefficients of connections between neurons are adjusted based on the input data. In other words, NN "learn" through examples (as children learn to recognize a particular subject, object, process or development through relevant examples) and have ability for generalization after use of training data. The artificial neurons (Figure 1), have a simple structure having similar functions as well as biological neurons. Body of the neuron is called node or unit.

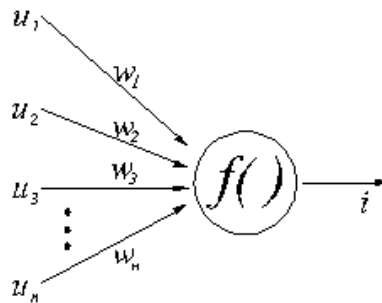


Figure 1: The artificial neurons

Where is:

$u_{1..n}$ – input data

$w_{1..n}$ – weight coefficients

$f()$ – activation function

i – output data

Architecture NM is a specific arrangement and connection of neurons in the grid. By architecture, NM vary according to the number of neural layers. Typically, each layer receives inputs from the previous layer and sends its output to next layer (Figure 2). The first layer is called input layer, the final is called output layer, the other layers are usually called hidden layers. One of the most common architecture of NM is the three layer network. The first layer (input) is the only layer that receives signals from the environment. The first layer transmits signals to the next layer (hidden layer) which processes the data and extracts the characteristics and patterns of the received signals. Data that are considered important are transferred to the output layer, the last layer of the network. On the output of neurons of the third layer we get the final results of processing. Complex NN may have more hidden layers, feedback loops and time delay elements, which are designed to enable more efficient separation of the essential properties or patterns from input level.

Learning NN comes down to learning from examples, which should be more as possible, that the network could behave precisely in a later exploitation. The learning process leads to the adjustment of synaptic weights.

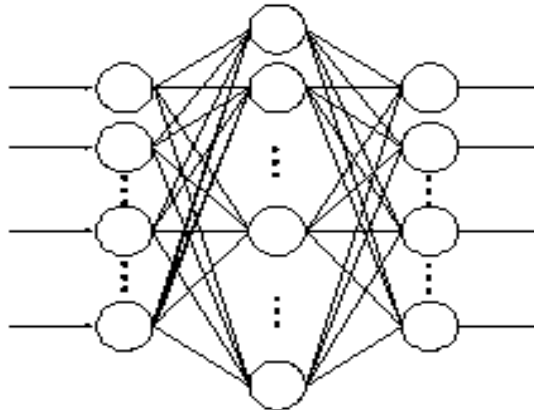


Figure 2.

When samples, that represent the network, does not lead to more changes in these coefficients, it is considered that the network is trained. There are a number of different implementation NN, and thus there is a lot of different classification. NN can be classified according to the number of layers, the type of connections between neurons, the type of training the neural network, according to the direction of propagation of information and type of data.

According to the number of layers, the NN can be divided into single- and multi-layered. Today, the most studied and applied multilayer NN, which in addition to the input and output layers containing neurons in the middle (hidden) layers.

GENETIC PROGRAMMING

Genetic algorithms are stochastic search techniques based on an analogy with natural processes. An analogy with the process of evolution in nature make base of the genetic algorithms. Genetic algorithms work with population of individuals who are often called strings or chromosomes. In the practical application of each string (chromosome) represents a solution to a particular problem. Each string consists of a finite number of elements, which are called genes. Genes are actually variables of the problem. Genes can take values from a finite set of values known as the alphabet. The process of finding a solution, by using a genetic algorithm, consists of several stages of evolution. Within each stage is formed a generation of solutions. From one generation to the next is crossed by application of genetic operators to the solutions of previous generation. To form a population of solutions of the next generation, it is necessary to apply three genetic operators: reproduction, crossover and mutation on previous generation.

Genetic programming is a branch of genetic algorithms (Poli, 2008). The main difference between genetic programming and genetic algorithms is how they present solutions. Genetic programming generates computer program as the solution in one of the programming languages, and genetic algorithms provide a solution in the form of a strings of numbers. Genetic programming uses four steps to solve the given problem:

- 1) generate the initial population of composition of functions which are randomly drawn and computer programs
- 2) execute each program in the population and determine its fitness by how good solve the problem
- 3) create a new population of programs

- make copy of the best existing program
 - creates new programs by mutations
 - creates new programs by crossbreeding (sexual reproduction)
- 4) the best computer program that appeared in any generation, the best solution to that moment is a result of genetic programming.

The algorithm of genetic programming is presented on Figure 3.

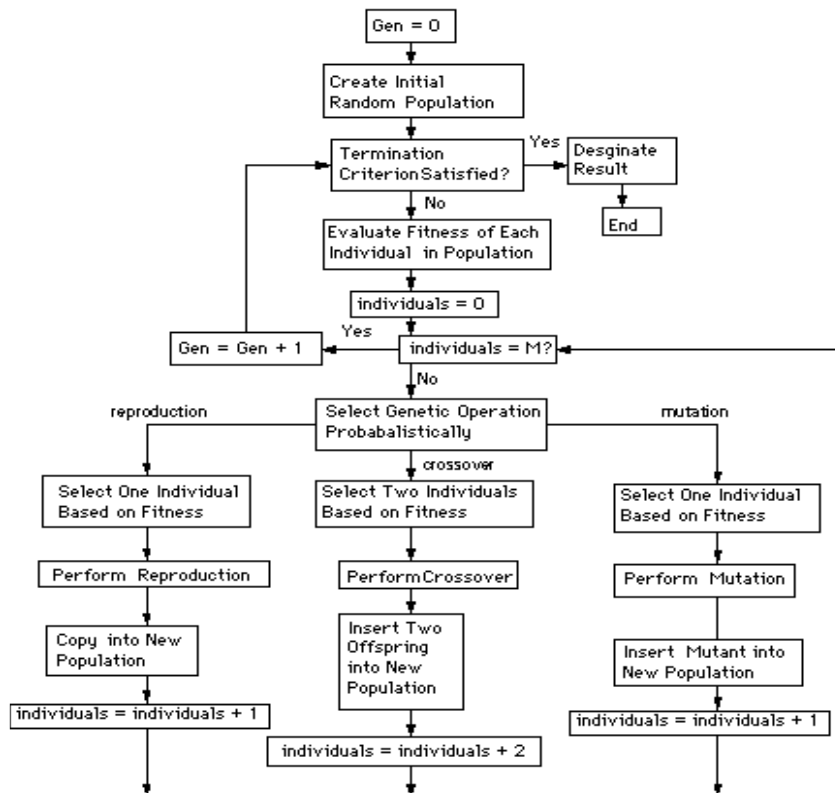


Figure 3: The algorithm of genetic programming

DEFINING AND SELECTION OF INPUT PARAMETERS

As already mentioned the aim of this study was to find certain dependencies between variables that are related to the number of employees and data which represent airline capacity for service. In addition it shall be ascertained whether the assumption that the use of artificial intelligence tools may provide quality and useful results, which could be used in the planning of the required number of employees.

Since, in this paper we are focused on number of employees in the airline, it is necessary that data, which represent capacity and capability of airline to provide services, be logically related to the number of employees, i.e. that number of employees depend on these quantities. Two parameters that can be correlated with the number of employees are number of aircrafts in the fleet and the number of available seat-kilometers.

Number of aircrafts in the fleet is an indicator of capacity and airline capability to provide services, but it alone is not sufficient to express the total capacity of the airline fleet. Number of aircraft in the fleet has a direct impact on the determining of part of the total number of employees which refers to the total number of crew and technical staff. This part of the employees is in direct proportion to the number of aircrafts, as each airline defines the number of crews that are rotated on a particular plane, on the other side flight crew is consist of three members (pilot, co-pilot and flight engineer). Also, each aircraft has a clearly defined number of cabin crew members, etc. The number of available seat-kilometers, is the relative parameter that clearly defines the airline capacity and capabilities to transport passengers with a fleet already determined (the number of aircrafts in the fleet, passenger cabin configuration, etc.) on a predetermined network. The assumption in this paper is that the number of aircrafts in the fleet and the number of available seat-kilometers are directly related to the number of employees in the airline. Data used in this paper are taken from 29 different airlines annual reports for year 2012. These data are shown in Table 1., where:

- Fleet (No) - the number of aircrafts in the fleet,
- Ask (000) - the number of available seat-kilometers in thousands and
- Noe - number of employees.

Table 1: Airlines annual reports for year 2012.

Airliner	Fleet (No)	AsK (000)	NoE (No)
Aeroflot	128	64880000	16418
Air Canada	351	124582188	24000
Air China	461	161382140	25269
ANA	226	56756000	32884
Austrian	74	23154000	6236
Avianca	147	36545000	18000
British Airways	273	158247000	43213
Cathay Pacific	138	129595000	21986
Croatia Airlines	13	2086000	1128
Delta	717	426728580	73561
Easy Jet	214	72182000	8446
El Al	24	20905000	5575
Emirates	169	200687000	33634
Ethiopian Airlines	48	22394000	6559
JAL	216	81189000	30882
Kenya Airways	34	13875000	4834
KLM	203	100727000	31189
Latam Airlines	311	31621000	22489
Lufthansa	627	259861000	116957

RyanAir	294	131750724	8438
SAS	184	36126000	14897
Singapore Airlines	100	113409700	13893
South African Airways	60	32423000	11044
TAP	71	35451000	7476
Thai Airways	101	79231000	25323
Turkish Airlines	202	96124000	15857
US Airways	433	163763100	30170
Virgin	94	39800000	8367
WestJet	100	40862528	7742

RESULTS

For the regression analysis were used professional software tools DataFit version 8.1.69 and CurveExpert version 1.4, while for the genetic programming we used Evolutionary Algorithm Program version 1.3 (Poli et al, 2008), (Una-May et al, 1992) and (Software, 1998) and NM Pythia - The Neural Network Designer version 1.02 (Software, 2000) Software package for genetic programming and NM are not "user-friendly" and for their use is required prior knowledge in these areas. Interface of these packages are shown in Figures 4 and 5.

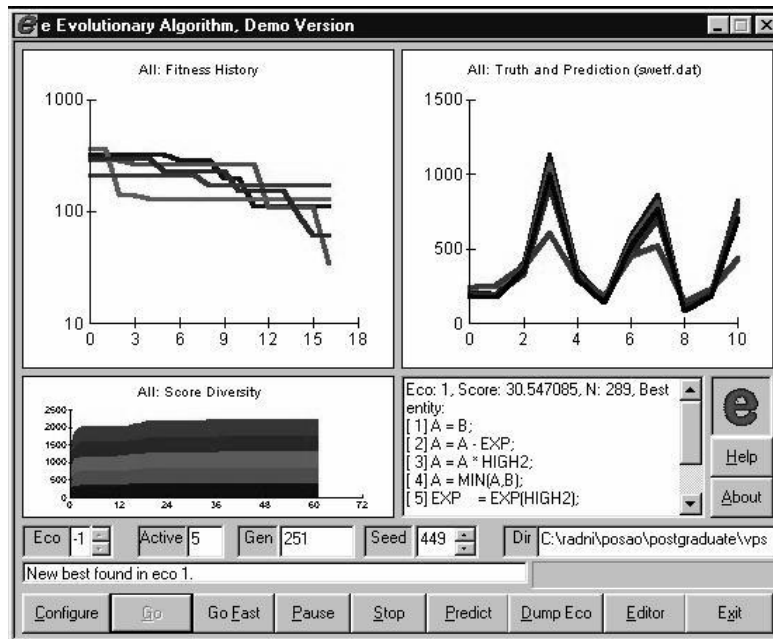


Figure 4: Evolutionary Algorithm Program

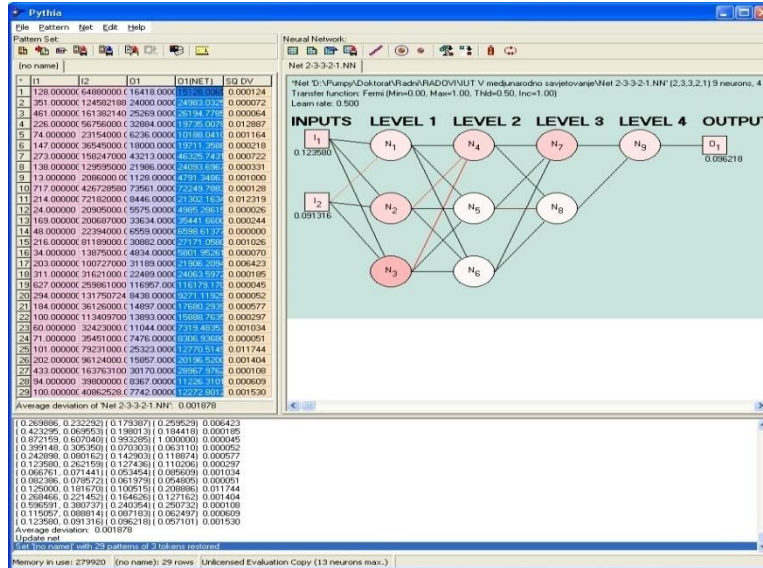


Figure 5: Software package NM Pythia

Learning results obtained by genetic programming are represented in the following pseudocode:

- [1] NoE = LOW1;
- [2] NoE = NoE * 6.21155;
- [3] NoE = NoE * 3.18256;
- [4] NoE = NoE * 3.18256;
- [5] NoE = NoE + (-1);
- [6] NoE = NoE + F;
- [7] TEMP2 = NoE;
- [8] COS = cos(TEMP2);
- [9] NoE = NoE * (-1);
- [10] NoE = NoE * LOW1;
- [11] NoE = NoE * COS;
- [12] NoE = MAX(NoE,9.81);
- [13] NoE = NoE - 61.2159;
- [14] NoE = NoE * 2.71828;
- [15] NoE = NoE - 61.2159;
- [17] NoE = NoE - 61.2159;
- [18] NoE = NoE - 61.2159;
- [19] NoE = NoE / F;
- [20] NoE = NoE * 6.21155;
- [21] NoE = NoE + 3.14159;

- [22] NoE = NoE + COS;
 [23] NoE = NoE * 3.18256;
 [24] NoE = NoE + HIGH2;
 [25] NoE = NoE - 61.2159;
 [26] NoE = NoE * COS;
 [27] NoE = NoE + LOW1;
 [28] NoE = MAX(NoE,61.2159);
 [29] NoE = NoE + (-1);
 [30] NoE = NoE - 3.18256;
 [31] NoE = NoE + LOW1;
 [32] NoE = NoE * 61.2159;

On Figure 6. we see that the NM consists of 1 input, 4 hidden and 1 output layer, ie. of 12 neurons in total. It has been shown that number of neurons larger than 12 in the NM does not affect on the final results of the analysis and the learning NM with larger number of neurons does not give a significant improvement compared to the results obtained by the learning network with 12 neurons.

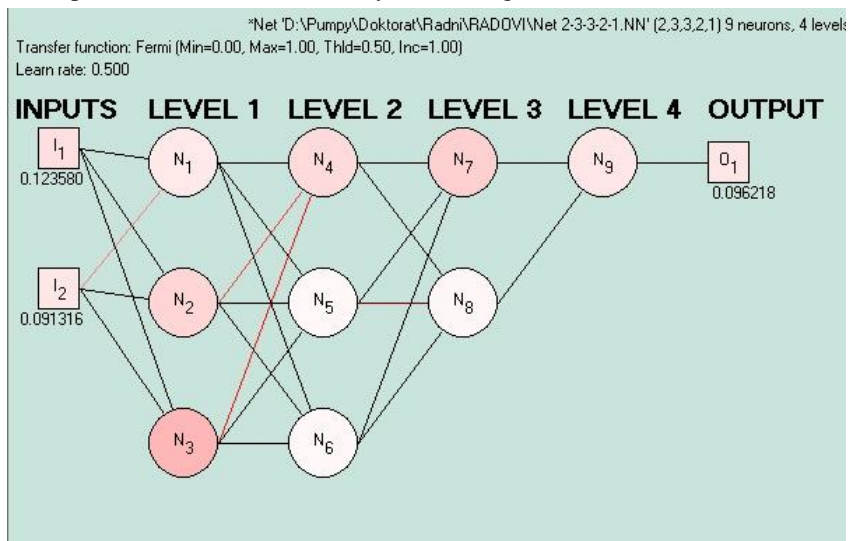


Figure 6: Result data obtained in relation to the actual data

Table 2. presents the results obtained by using genetic programming, NM and regression analysis. On Figure 6. is shown distribution of result data obtained in relation to the actual data. From the diagram it can be seen that the results obtained by NM have the best fit with the actual data.

Table 2: Results obtained by using genetic programming, NM and regression analysis.

Inputs		Truth	GP	NM	Reg
Fleet (No)	AsK (000)	NoE	NoE	NoE	NoE
128	64880000	16418	16421	15128	18488
351	124582188	24000	24978	24983	25843
461	161382140	25269	31712	26195	31116
226	56756000	32884	29855	19735	20722
74	23154000	6236	8021	10188	8123
147	36545000	18000	18950	19711	14074
273	158247000	43213	42730	46326	24354
138	129595000	21986	11939	24094	18571
13	2086000	1128	4287	4791	371
717	426728580	73561	51727	72250	73561
214	72182000	8446	16592	21302	22006
24	20905000	5575	4961	4985	5827
169	200687000	33634	25002	35442	36572
48	22394000	6559	6430	6599	7044
216	81189000	30882	33489	27171	22512
34	13875000	4834	5573	5802	4194
203	100727000	31189	28054	21906	21941
311	31621000	22489	22530	24064	18319
627	259861000	116957	106200	116179	116685
294	131750724	8438	21489	9271	23739
184	36126000	14897	14755	17680	15209
100	113409700	13893	10665	15889	17948
60	32423000	11044	7164	7319	10139
71	35451000	7476	7838	8307	11268
101	79231000	25323	12107	12771	18606
202	96124000	15857	15857	20197	22045
433	163763100	30170	29998	28968	30603
94	39800000	8367	9246	11226	13072
100	40862528	7742	10665	12273	13514

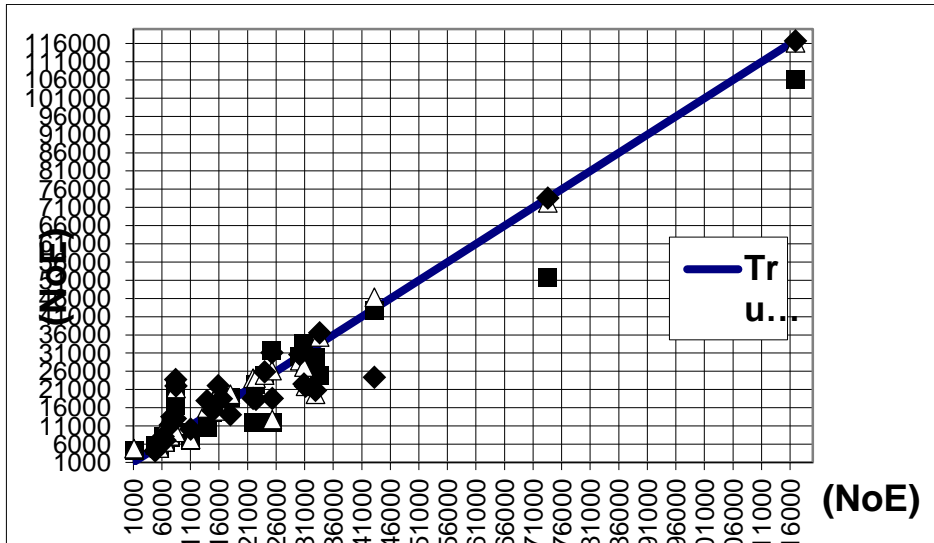


Figure 7: Results obtained by NM have the best fit with the actual data.

Table 3. presents the correlation coefficients (R) and standard error (S) obtained by regression analysis (REG), genetic programming (GP) and the application of artificial neural networks (NM). It may be noted that results obtained by NM and genetic programming are better, i.e. correlation coefficients are greater than on results obtained by regression analysis. These results confirm that correlation between the parameters exists and it is significant.

Table 3: Correlation coefficients (R) and standard error (S) obtained by regression analysis (REG), genetic programming (GP) and the application of artificial neural networks (NM).

		GP	NM	Reg
NoE	R	0.9644	0.976	0.9552
	S	5412.28	5035.67	6748.71

CONCLUSION

Based on the results obtained using the NN and genetic programming can be concluded that it is reasonable to presume that the tools of artificial intelligence can be applied even for complex tasks such as planning and management of human resources in the airline. Planning is a management function, which plays a very important role in modern organizations (Kastratović, 2008). Obtained results can be used for planning the number of employees, i.e. planning the necessary financial investments in human resources, and may also be useful for a preliminary analysis in airlines which decide to perform restructuring plan or increase / decrease the number of operations. Due to a high degree of competitiveness and the globalization of the modern market, each business segment must be thoroughly planned (Kastratović et al. 2016). Since airlines regularly publish business reports, management can perform analysis and comparison results achieved by other airlines with similar fleet and network on which they provides services of transport of passengers. By using NN and genetic programming in the analysis it is possible to reached dependencies that are not obvious at first glance. This is particularly important in a competitive and different business models.

For quality results, and for good training of NN and genetic algorithms we need great source of high quality data. Besides this it is important to have a good knowledge of theoretical and practical dependences of input and output parameters, as it is very shortens the processing of data and avoids idling because of possible poor connection of some parameters or even their incongruity. It is also shown that NN and genetic programming provides better results than conventional models, which are based on statistical and regression analysis. The real quality of the results should be shown on large number of processed data. Work is the result of project no. FS III1 / 0318-0319 / 208-2019.

REFERENCES

- [1] Carsten P., Thorsteinn, R. (1992). "An introduction to artificial neural networks", Proc. 1991 CERN Summer School of Computing, CERN Yellow Report 92-02, 113-170 (1992.)
- [2] David K. (2007). "A Brief Introduction to Neural Networks", 2007, available at <http://www.dkriesel.com>
- [3] Kastratović E. (2008).,; Uvod u menadžment. Beograd: Fakultet za menadžment u sportu Univerziteta "Braća Karić", dopunjeno izdanje, 2008. ISBN 978-86-86197-24-5. [COBISS.SR-ID 154583564]
- [4] Kastratovic, E., Kalicanin, M., Kalicanin, Z.(2017). "Features of Cash Flow Compared to Profit", International Review, (2017. No.3-4), http://www.vspep.edu.rs/img/downsekcija/2018/02/ejournal_international_review_no_3_4_2017.pdf
- [5] Petrović, D., Mirosavljević, P.(2000). "Determining of Geometry of Transport Aircraft Using Genetic Programming Tools", paper accepted for exposing and publishing on SYM-OP-IS 2000, XXVII Yugoslav Symposium of Operational Research, Vrnjačka Banja, 10-13 Oct. 2000.
- [6] Petrović, D.,(2001). "Estimating Transport Aircraft Stalling Speed by Using Artificial Neural Networks", paper accepted for exposing and publishing on SYM-OP-IS 2001, XXVIII Yugoslav Symposium of Operational Research, Belgrade, 02-05 Oct. 2001, Proceedings, ppg. 637-640
- [7] Petrović, D. (2001). "Preliminary Definition of Flight Properties for Transport Aircraft with Turbo-Jet Powerplant by Using of Genetic Programming Tools", Master thesis, Traffic and transport faculty, University of Belgrade, Belgrade, 2001.
- [8] Puharić, M., Lučanin V.,Petrović D. (2013). "Design and maintenance of technical systems with emphasis on the maintenance of rail vehicles and aircraft", Monography, Institut Goša, Belgrade 2013.
- [9] Pythia – The Neural Network Designer, Version 1.02, Runtime Software, 2000.
- [10] Riccardo, P., Langdon, W., McPhee, n. (2008). "A field guide to genetic programming", Published via <http://lulu.com> and freely available at <http://www.gp-field-guide.org.uk>, 2008. (With contributions by John R. Koza)
- [11] O'Reilly, U., Oppacher, f. (1992). "An Experimental Perspective on Genetic Programming", Parallel Problem Solving from Nature II, Brussels, Belgium, Oct, 1992.
- [12] "User's Manual, e Evolutionary Algorithm Program for MS-DOS, Windows 3.1, Windows 95, Windows NT, UNIX, and OS/2 Operating Systems", System Dynamics International, Inc., V980303, 1998.

Article history:

- Received 20 September 2018
- Accepted 25 November 2018