IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 14, Issue 4 Ver. II (Jul - Aug 2018), PP 50-57
www.iosrjournals.org

Examining the Limited Rational Decisions of Neural Networks

Tzotzis Christos[1], Kalampakas Argyrios[2], Makris C. Georgios[3]
[1]*Postgraduate student of Inter-Faculty Master Program on Complex Systems and Networks of Aristotle University of Thessaloniki.*
[2] *Postgraduate student of Inter-Faculty Master Program on Complex Systems and Networks of Aristotle University of Thessaloniki.*
[3]*Department of Mathematics, Aristotle University of Thessaloniki, Greece*

Abstract: *In the present work 12 different artificial neural networks are constructed to control the limitations of rational decisions taken by them. Forecasts are being studied in relation to the number of repetitions of education combined with the time of training. Artificial neural networks are used to predict the average property price.*
Keywords -*Neural Networks, Bounded Rationality, MSE, Limited Rational Decisions*

Date of Submission: 17-07-2018 Date of acceptance: 02-08-2018

I. Introduction

The purpose of this paper is to construct neural network techniques (NNTs), which will test the limitations of their rational decisions. The objectives of the thesis are to check the following claims:

- Increasing the memory and the speed of the NNs, the prediction results are improved.
- There is a limit to increasing memory and speed so that after this the NN gives almost the same results.

For the training of the NNs, data from the Housing Values in Suburbs of Boston experiment was derived from the MASS library of Statistical Packet R.

It describes the way in which man makes rational decisions and their limitations. The general way of operating the neural networks, their structure and the most common method of training will be presented.

Bounded Rationality

The rational decision is the best choice in terms of resource and condition data. It makes it possible to choose between all possible alternatives. The rational ability requires a thorough knowledge of the consequences that will follow, after each choice. People, in order to cope with the complexity of problems, adopt different decision-making strategies. These strategies allow them to "simplify" the problem and focus on its more obvious features [1].

When making a decision there is an interaction between the man and the environment. When the future is uncertain, the information is incomplete, the time and the resources are insufficient, we are leading to decisions of bounded rationality. People try to be rational, but their decisions are limited by their cognitive abilities, the adequate information and the limitations of the problem. Thus, they do not thoroughly examine all the elements but they choose as the best solution, the one that seems to serve the needs and be "satisfactory" [2], [3], [4].

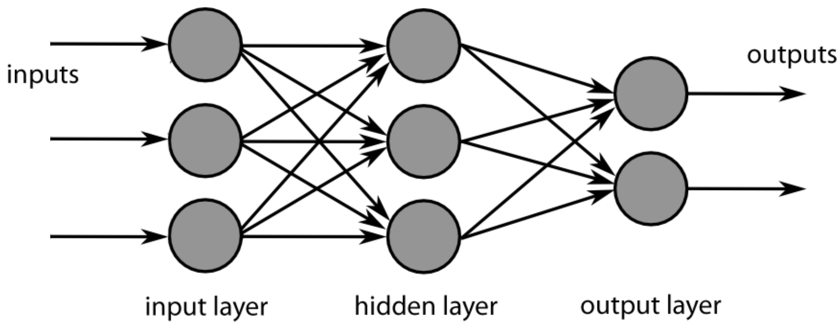
Herbert A. Simon [5], [6], [7] stressed that cognitive processes that trigger the bounded rationality are not intended to maximize the profit or the utility, but to find a satisfactory solution. Decision-making aims to select satisfactory alternatives, and only in exceptional cases to discover and select the best alternatives. An alternative solution is considered to be the best if it is superior to all the others, according to a single set of criteria used to compare all available alternatives.

Neural Networks

An artificial neural network is a computational model inspired by the way the brain's biological networks process the information. The basic computing unit of the neural network is the neuron. It receives information from variables or an external source and calculates a result. Each imported information has a weight assigned to it according to its importance in relation to the others. The neuron applies an activation function f to the input information [8].

The simplest type of neural network is the "feedforward neural network". It contains multiple neurons ranked in levels.

Figure 1: *Artificial neural network*



Input Level: It contains the neurons that accept the values of the variables.
Hidden levels: They perform calculations and transfer information from the first level to the last one.
Output Level: It contains the neurons where the results are extracted [9].
The most well-known method of training a neural network is the "backpropagation" method. Initially all the weights are set at random. For each data packet given in the training algorithm, the neural network is triggered and an output is given. This result compares to the desired one and their deviation is advanced to the previous level where the weights are adjusted. This process is repeated until the deviation is less than a predetermined threshold [10], [11].

II. Construction of Neural Network

The R language was used to conduct the experiment. The data from the Housing Values in Suburbs of Boston experiment come from the MASS library [12].

The construction of the technical neural network and its training was done with the help of the neuralnet package: Training of Neural Networks [13].

Data

The data which collected consists of 506 lines (values) and 14 columns (variables) and relates to property prices in the suburbs of Boston. The following is a description of the variables.

X_1: Crime rate
X_2: The proportion of habitable land over 25,000 m².
X_3: The percentage of non-retail businesses
X_4: If it is a riverside
X_5: Concentration of nitrogen oxide
X_6: Average number of rooms per house
X_7: Ownership rate for buildings before 1940
X_8: Average distances in 5 employment centers
X_9: accessibility index on a highway
X_{10}: Real estate tax per 10000 $
X_{11}: Proportion of pupils and teachers
X_{12}: percentage of African Americans in the city
X_{13}: Living standard percentage
Y_1: average house price per 1000 USD (predictive variable)

The data were divided into two random parts (not continuous) in a proportion of 75% -25%, where the 75% was used at the training stage and the 25% at the forecasting stage. Then normalization was performed with the min-max method in the data so that the neural network will give as better results.

Training and forecasting process

The neural network simulated the linear regression model $\alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_{13} X_{13} = Y_1$. The variables X_1 to X_{13} were used as input neurons and the variable Y_1 as an exit neuron. The training of the neural network is a process that has many parameters. The threshold for the rate of change of error, the maximum number of steps for a repetition, the number of repetitions of education and the learning rate.

All parameters in the default values were kept constant in the build-in NNs and only the number of iterations changed. In particular, the maximum error rate should be less than 0.01 and the training steps should be at most 100000.

12 different technical neural networks were built.

During the training of each neural network, the training time and the size it occupies in memory are stored. These two values were increasing as the number of iterations was increased. It was thought that the memory which required by the neural network corresponds to the experiences of a human being [14].

At the predictive stage, the only parameter that existed is the number of iterations. Also from this stage was saved the training time, which was considered as the available time that a person has to make a decision.

The first neural network was trained with 10 reps and predicted with 10 repetitions. In each of the next 4 neural networks, the number of iterations of education doubled, while the number of iterations of the forecast remained constant. In the next 4 neural networks the number of iterations of the forecast was doubled, while the number of iterations of the training remained constant. Having the same number of iterations of training and forecasting, in the next 3 neural networks both repetitions were doubled. There were a total of 12 neural networks with a different combination of repetition numbers for training and predictions.

III. Results

Figure 2: *Predictive values to the actual values for the 1st neural network.*

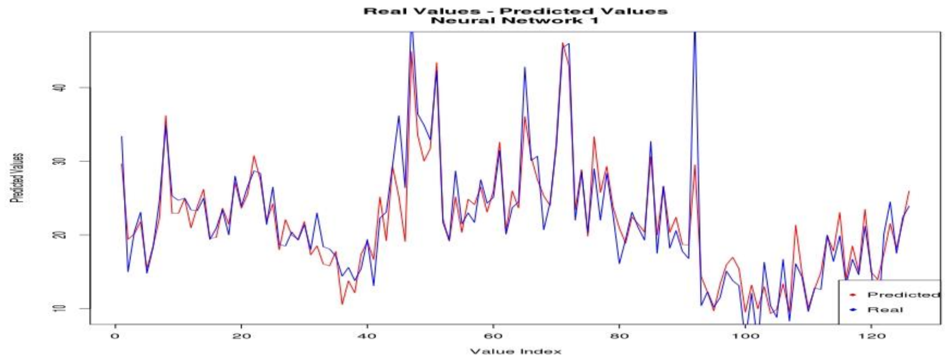


Figure 3: *Predictive values to the actual values for the 4th neural network.*

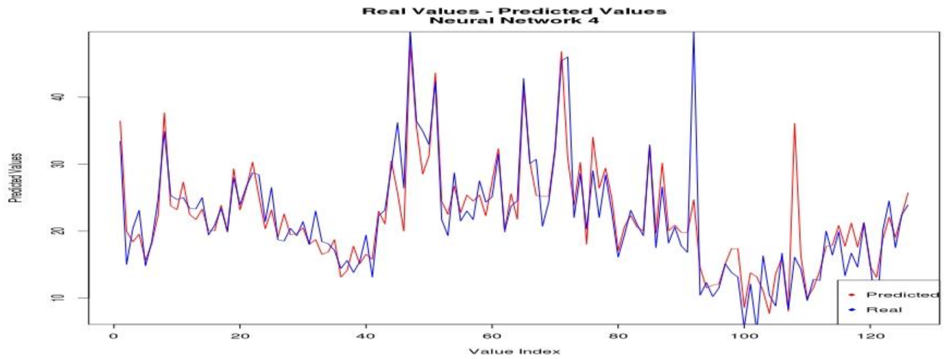


Figure 4: *Predictive values to the actual values for the 8th neural network.*

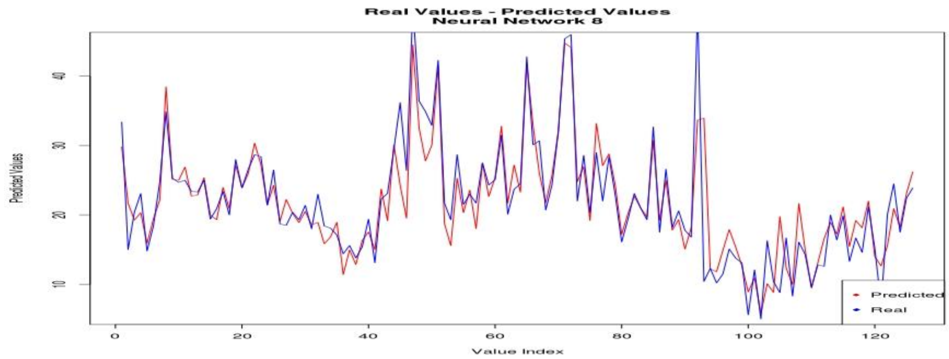
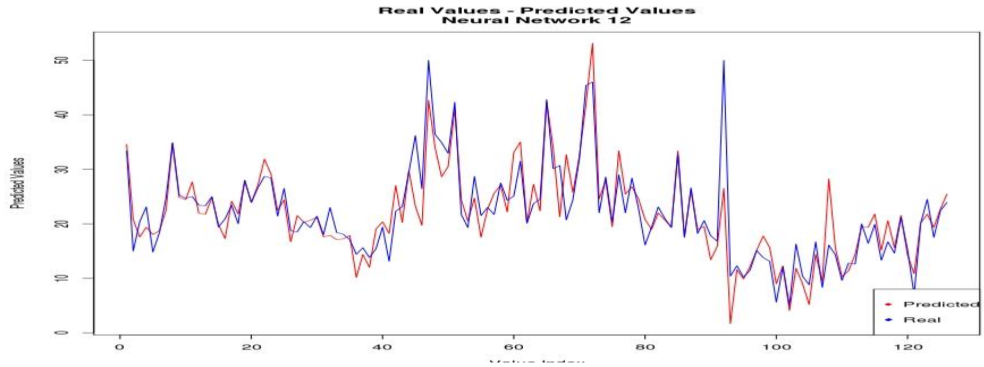


Figure 5: *Predictive values to the actual values for the 12th neural network.*



The previous figure show the predictive values to the actual values. Because the prediction was made for 25% of the data and not for a single value, the mean squared error (MSE) was used as the general error indicator. MSE does not decrease as the neural memory increases, nor does the speed of training increase.

Figure 6: *The MSE in each neural network.*

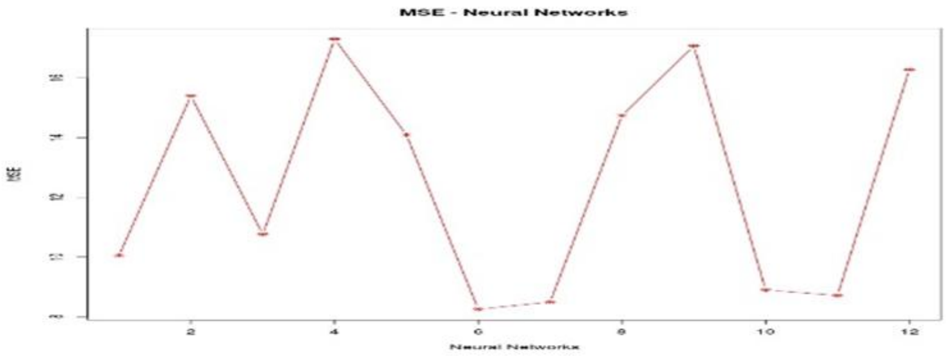


Figure 7: *The MSE compared with the neural network memory.*

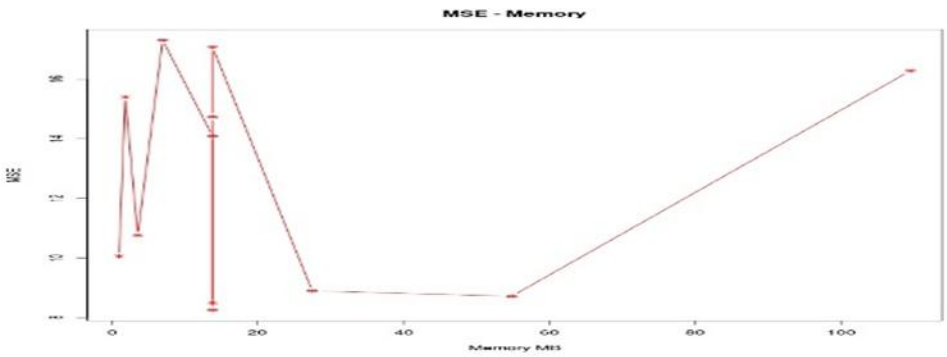
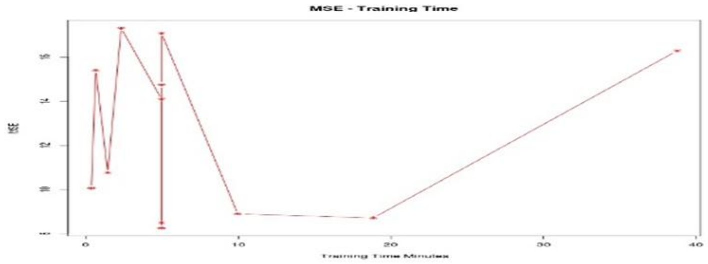


Figure 8: *The MSE compared with the training time of the neural network.*



The MSE does not decrease during the progression of the neural network, but it is observed that the MSE has the lowest value for the 6th neural, which had a ratio 8: 1 number of repetitions of training with a number of repetitions of prediction.

At this point, it was observed that the network has "memorized" the results of the education phase but has not learned to generalize them correctly on the new entry data. This phenomenon is called overfitting and brings about a relatively small error on the outputs results from the training of the network but also a fairly significant error in the outputs occurs when new data are given to the network [15], [16]. Due to overfitting and observation for the ratio of the 6th neural, the 8: 1 ratio was chosen and the neural network developed from the beginning.

Figure 9: The *actual values compared with the predicted values for the 1st neural network.*

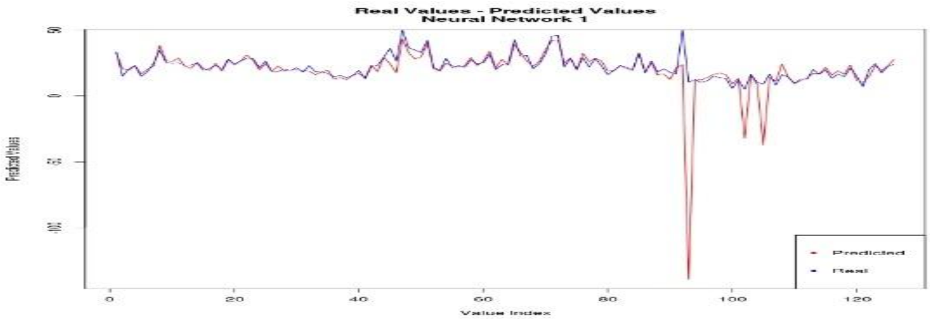


Figure 10: *The actual values compared with the predicted values for the 3rd neural network.*

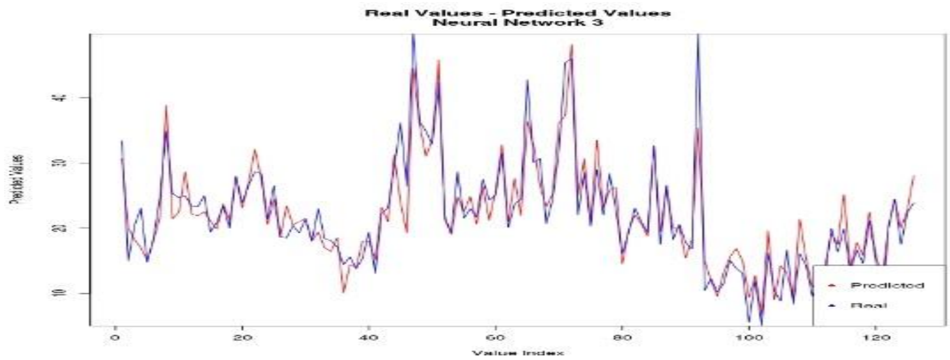
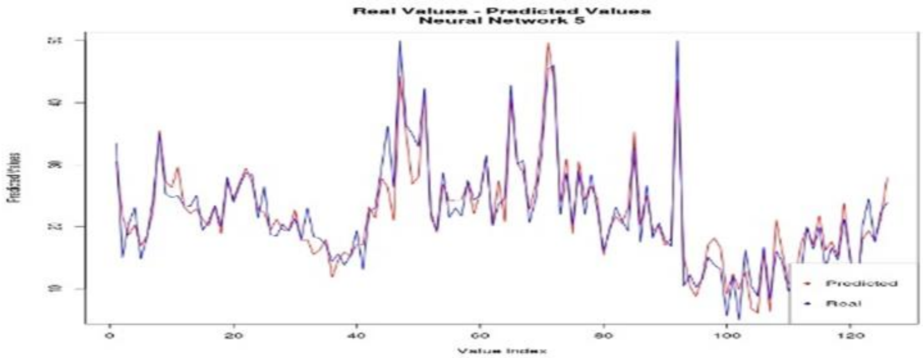
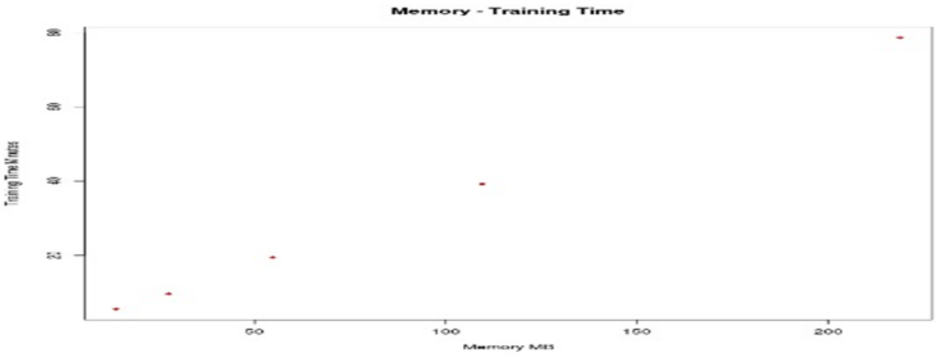


Figure 11: *The actual values compared with the predicted values for the 5th neural network.*



There is an adjustment of the forecast to the actual value outside the extreme values.

Figure 12: *The memory of the neural network compared with the training time of it.*



Note that there is a proportion of training and memory time.

Figure 13: *The MSE of the neural network compared with the memory of it.*

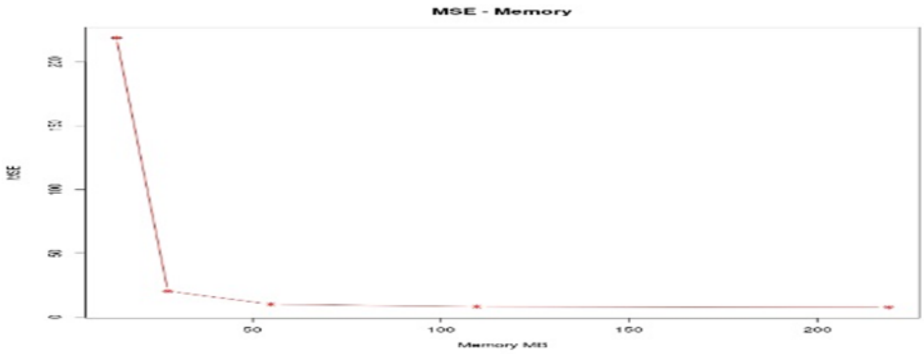


Figure 14: *The MSE of the neural network compared with the training time of it.*

Figure 15: *The MSE of the neural networks.*

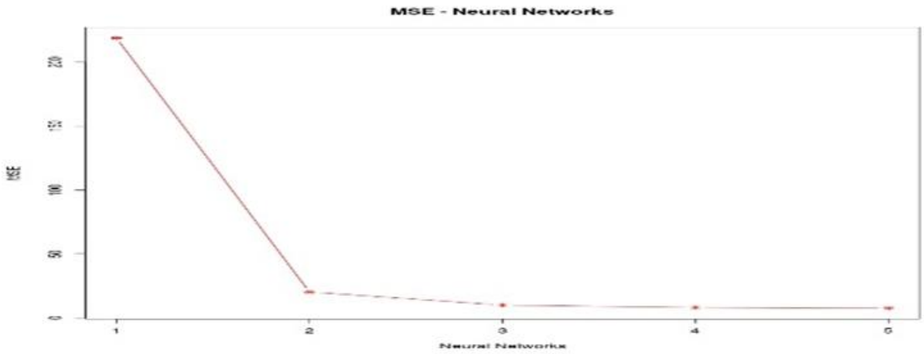
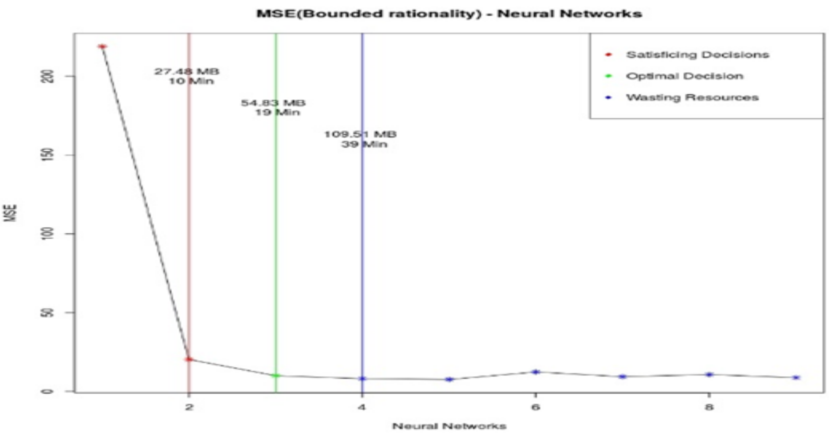


Figure 16: *The MSE as bounded rationality index for each neural network.*



The MSE decreases as the memory of the neural network increases, training time, but also during the evolution of the neural network.

As Bounded Rationality was considered the MSE index. When the MSE is equal to zero, the neural network will make a rational decision. Specifically, the third neural network required 54.83 MB of memory and 19Min of training time, with 640 training reps and 80 predictions repeats, to reach a rational decision.

IV. CONCLUSION

The conclusions drawn from the above constructions and measurements are:

- As the memory and the speed are increasing, without compromising the ratio, the results are improved, so the limitation which imposed by the limited rationality applies.

- With a limited memory and time, a decision of limited rationality is taken, and as the memory and the time are increasing, the answer is improved.
- It is noted that in the model that was implemented, after the third neural network, the MSE remains at the same level as if the neural network resources increased. It is believed that the rational decision has been taken from the third neural network and all subsequent neural networks continue to make the same decision, "wasting" more resources. As for the first two neural networks, they are the neural networks that have taken a satisficing decision.

References

[1]. Bayer, R. C., Renner, E., &Sausgruber, R. (2009). *Confusion and reinforcement learning in experimental public goods games*. NRN working papers 2009-22, The Austrian Center for Labor Economics and the Analysis of the Welfare State, Johannes Kepler University Linz, Austria.

[2]. https://www.tutorialspoint.com/management_concepts/decision_making_process.htm (date of access 25-07-2018)

[3]. Colman, Andrew (2006). A Dictionary of Psychology. New York: Oxford University Press. p. 670. ISBN 0-19-861035-1.

[4]. Gigerenzer, Gerd; Selten, Reinhard (2002). *Bounded Rationality: The Adaptive Toolbox*. MIT Press. ISBN 0-262-57164-1

[5]. Herbert A. Simon, 1972, *Theories of bounded rationality*, North-Holland Publishing Company

[6]. Herbert A. Simon, 1982, *Models of Bounded Rationality, Vols. 1, 2 (1982), 3 (1997)*, MIT Press

[7]. Herbert A. Simon, *A Behavioral Model of Rational Choice*, The Quarterly Journal of Economics, Vol. 69, No. 1 (Feb., 1955), pp. 99-118

[8]. D. Graupe, "Principles of Artificial Neural Networks.3rd Edition", World Scientific Publishers, 2013

[9]. Auer, Peter; Harald Burgsteiner; Wolfgang Maass (2008). "A learning rule for very simple universal approximators consisting of a single layer of perceptrons". *Neural Networks*. **21** (5): 786–795. doi:10.1016/j.neunet.2007.12.036.

[10]. Nielsen, Michael A. (2015). "Chapter 6". *Neural Networks and Deep Learning*.

[11]. A. W. Harley, "An Interactive Node-Link Visualization of Convolutional Neural Networks," in ISVC, pages 867-877, 2015

[12]. https://cran.r-project.org/web/packages/MASS/index.html(date of access 25-07-2018)

[13]. https://cran.r-project.org/web/packages/neuralnet/index.html (date of access 25-07-2018)

[14]. https://www.r-bloggers.com/fitting-a-neural-network-in-r-neuralnet-package/ (date of access 25-07-2018)

[15]. Burnham, K. P.; Anderson, D. R. (2002), *Model Selection and Multimodel Inference* (2nd ed.), Springer-Verlag

[16]. https://papers.nips.cc/paper/1895-overfitting-in-neural-nets-backpropagation-conjugate-gradient-and-early-stopping.pdf (date of access 25-07-2018)

Tzotzis Christos " Examining the Limited Rational Decisions of Neural Networks" IOSR Journal of Mathematics (IOSR-JM) 14.4 (2018) PP: 50-57.