Journal of Hyperstructures 9 *(Spec. 3rd CSC2019)* (*2020*), *1-10.*
ISSN: 2322-1666 print/2251-8436 online

COMPARING THE NEURAL NETWORK WITH PATH ANALYSIS IN FITTING REGRESSION MODELS

FERESHTEH ARAD* AND AYYUB SHEIKHI

Abstract. The purpose of this study was to compare a neural network with the path analysis in fitting regression models. The conceptual model of path analysis according to the studied data includes a dependent variable, two independent variables and one mediating variable. The neural network conceptual model is considered with three layers (input, hidden, output) and the hidden layers have two nodes. The study asked 474 people about their education, beginning salary, previous experience and their current salaries. The data divided into the train and test groups at the rate of %60 and %40. The criterion for comparing the two methods is RMSE. The results of the analysis showed that both models are over fitted and the RMSE train and test of neural network are less different from the path analysis. Therefore, in this dataset, it can be said that the neural network performs better than the path analysis.

Key Words: Neural network, Path analysis, Regression, RMSE.
2010 Mathematics Subject Classification: Primary: 62M10; Secondary: 62M45.

1. Introduction

For the first time a biologist, named Seville Wright (1918), used the direct and indirect effects analysis on a set of variables and introduced a general case of the traditional regression namely of path analysis. Much of his work was on models that included manifest variables. In 1950,

Received: 19 November 2019, Accepted 01 August 2020. Communicated by Behrouz Fathi-Vajargah;
*Address correspondence to Fereshteh Arad; E-mail:fereshteh_arad@yahoo.com

economists became aware of Wright's work and were surprised to find that he was ahead of their estimates in supply and demand, and used his ideas in econometrics [1]. However, the path analysis method for modeling relationships between variables has a number of assumptions and limitations. These limitations include the existence of a linear relationship between the predictor variables with the dependent variable, the independence of the errors terms of the independent variables, the normality of the data, and absence the of multicollinearity. Given the limitations mentioned for the path analysis method, there is a need for methods that face fewer constraints in this regard. In the meantime, the neural network can be one of the most appropriate methods. The first scientific application of neural networks was introduced in the late 50s of 20th century, when Rosenblatt et al 1958, built networks that were able to identify patterns [5]. The development of neural networks continued until the 70s of the 20th century. the idea of using a stochastic mechanism to explain the operation of a broad class of recursive networks that can be used to store information was introduced by the American physicist John Hopfield in 1982 [7]. The second important idea that became the key to the development of neural networks in the 80s of the 20th century is the "back-propagation" algorithm was proposed by David Rommelhart and James McClend in 1986 [3]. Over time, many people have been working on a variety of neural networks [4, 5, 6]. In addition, compared to path analysis, neural networks are powerful tools for nonlinear modeling. Neural networks are a new way of modeling that, due to their intelligent and flexible structure, are competing with conventional statistical modeling and are advancing in both theoretical and applied contexts. These methods do not impose any initial assumptions on the distribution of data, and are especially valuable when the functional relationship between independent and dependent variables is unclear, another advantage of neural networks is that they can handle big data and data is processed in the network implicitly, even if some parts of the network layers are removed or not functioning properly, correct answers are still possible. The generalizability of the network allows the model to respond well to a new observation. In this study, we compared two models of path analysis and neural network fit regression models with an example.

2. Preparation

2-1 Path analysis

In multiple regression, independent variables explain dependent variable. The question that may arise to the researcher is whether their own independent variables affect each other. In fact, researchers sometimes try to test a set of relationships, some of which test the relationship between the dependent variable and the independent variables and other examine the relationship between the independent variables themselves. So there is a set of relationships that are introduced and tested in a model. To simultaneously investigate these tests, a branch of statistics is used in the name of path analysis, which has been highly regarded in the last few decades. In fact, as we will see later, path analysis uses the standardized regression as well as the standardized beta so that the relationship between a set of variables can be shown in the form of causal relationships. To be able to do the path analysis using regression we should first be able to convert the ordinary regression equation to the standard regression equation. Denoting the standard deviation of the dependent variable and the ith independent variable as S_y and S_i respectively, it can be show that

$$\beta_i^* = \beta_i \times S_i/S_y \tag{2.1}$$

By having standard path coefficients, we can convert the regression into the path analysis. In addition, using path analysis, we can show the correlation between variables as direct effect, indirect effect and total effect. The question that might happen is that if we want to establish routes between the independent variables, which of them should we consider the dependent variable? In fact, we can perform the path analysis using regression. For this purpose, first, multiple regression can be carried out between the original dependent variable and the independent variables. Then, to examine the relation between the independent variables, among them we consider a variable as the new dependent variable, which has the highest correlation with the initial dependent variable. To identify that variable, we use the standardized beta coefficient, every variable that had the highest coefficient, is considered the second dependent variable and we conduct multiple regression between it and the other independent variables. Next step, the variable that has the most relation with the second dependent variable, we consider as the new dependent variable, we will continue to do so.

2-2 Neural network

Neurons are the constituent elements of the brain and act alone like a rational processing unit. The human brain consists of thousands of billions of neurons and each neuron is roughly connected to thousands of other Nero. Scientists believe learning takes place in the brain when connecting one cell to another at the synapse site is corrected. That is, Neurotransmitters with more ease can be released into the synaptic cleft, this makes the larger gates open to the opposite dendrites. This will allow the cell to be activated with less stimulation. The ability to shift the connection between cells is one of the most important characteristics of neural network patterns. This strengthening of connections is a learning type the so - called incentive learning. In this way of learning, trial and error are used to determine how much neurons are connected. That is, learning occurs in a way that, in order to do anything, to understand the cause and effect, to process the images received from the eye and etc. There is a unique neuron structure and this neuron structure will work best. In other words, data on everything must pass through a particular neuron. So, with the arrival of any kind of data, the neurons whose connections are amplified to process the data are more likely to be activated. So the activation of the first neurons, the data will be put forward in a pre-reinforced path. In general, a neural network has the following properties:

1) Method for calculation is made based on the interconnected link of several processing units.

2) The network consists of an arbitrary number of cells, nodes, or units, which link the input set to the output.

3) Artificial neural network is a practical method for learning various functions such as functions with real values, functions with discrete values and functions with vector values.

4) Neural network learning has been immune to educational faults and such networks have been successfully applied to issues such as speech recognition, recognition and interpretation of robot images and learning.

The neural network elements are:

Input and output: Numbers in format one or more variables, it forms the inputs of a neural network. These inputs are transformed into one or more output variables after performing specific analysis and processing. The inputs play the role of the independent variable and the outputs play the role of the dependent variable.

Neurons: the most important part is the nervous system of neurons. That are divided into three categories of input, output, and hidden neurons, in the form of, the input layer, the outer layer and the outer layer are placed. The neurons or input units have the responsibility to receive input data. Intermediate layer and output consist of data processing units that in these units, algebraic operations will be performed on the input data and their result is sent in the form of a new entry into other units in the next layers.
Weights: The input various variables of the network have different value, that the weights are allocated to them. These weights that are considered before the output layer and the hidden layer are generated by random number method and corrected using the grid.
In general, the neural network formula of several layers is as follows:

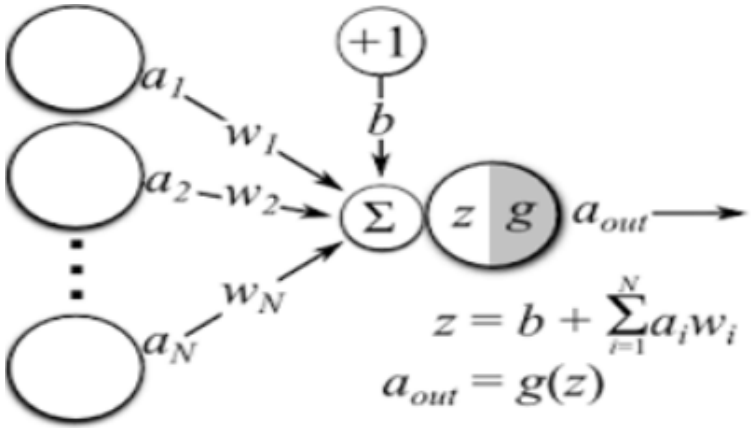

FIGURE 1. Neural network components

$$a_k = g_k(b_k + \sum_j g_j(b_j + \sum_i a_i w_{ij}) w_{jk}) \tag{2.2}$$

where in this formula a_k output, a_i inputs and w_{ij}, w_{jk} the weights are in each layer.
The advantages of neural network can be referred to as [2]:
1) The power of learning and adaptation: Neural networks can learn according to the responses in the subject environment that how to respond to the inputs and keep the learning in its memory.
2) The power of generalization: Neural networks can extend the results to similar cases after learning and adaptation.

3) Background information processing: Knowledge of any subject, it is expressed by the highly structured and active state of a neural network. Any neuron in the network is prone to affect the overall activity of other neurons. As a result, background information is naturally distributed by a neural network.
4) Error tolerance: The performance of a neural network will optimal if adverse conditions occur. In principle, a neural network, in the event of a problem, suffers a relatively favorable performance failure, not a catastrophic failure.
5) Low energy consumption: The neural network consumes little energy, due to parallel information processing and information retention.
Also, the model is divided into three groups in terms of fit accuracy of test and error models. The first group is high fit, it is so complex that it works great for experimental data, but that complexity it very bad for test data. The second group is low fit, it is so simple that it gives a very bad answer to the experimental data and this simplicity it perfect for the test data. And the third group is exactly fit, neither the high fit complexity nor the simplicity of the low fit, and this feature makes their accuracy in the experimental data as much as the test data.

3. CONCEPTUAL MODEL AND MODEL FIT

3-1 Model fitting through neural network

As mentioned before, we divided data set to the two train and test subsets at the rate of %60 (284 persons) and %40(190 persons). Then, we perform this neural network using R and codes are available upon request.Figure 2, shows this neural network, its weight, the estimating error and numbers of iterations. After fitting the model RMSE value listed in.

$$V_1 = g_1(b_1 + \sum_j g_j(b_j + \sum_{i=2}^{i=4} V_i w_{ij}) w_{j1}) \tag{3.1}$$

3-2 Fitting the model through the classical statistics (path analysis)

First, considering variable V_1(salary) as dependent variable, V_2(education), V_3(previous experience), V_4(beginning salary) as independent variables we have done a mult reg (Figure 3).
According to Table 1 because the variable V_4(beginning salary) had the highest impact on the salary variable we choose it as a new dependent variable. We then conducted the path analysis with AMOS

software with two independent variables(V_2(education)), V_3(previous experience)), a mediator variable (beginning salary V_4) and a dependent variable (V_1(salary)). All variable of this model are either manifest or observed. So we put them in a rectangle, error variables not visible so they are located in the circle by convention of structural equation models. The other arrows, which are one-sided, show the cause and effect relationship between the variables (Figure 4).
In the other hand, with %60 of the data, with 284 people, we trained a neural network and we tested it with another %40, and the RMSE value is given in the last column of Table 2.

According to the output of Table 2, the RMSEA for both methods is divided into train RMSEA and test RMSEA. The train RMSEA of path analysis is 0.603 and test RMSEA path analysis is 0.625. Also train RMSEA neural network equals 0.55 and test RMSEA neural network equals 0.55. The results show that both models are over fitted. In other words, the mean square error of the test is higher than the squared mean of the train. So, since the error between RMSEA test and train path analysis model are more, then it can be said that neural network method has priority over path analysis.

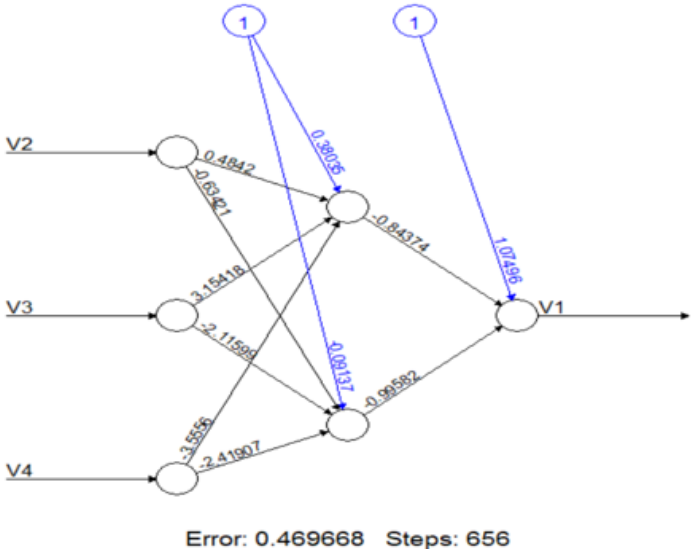

FIGURE 2. Neural network model

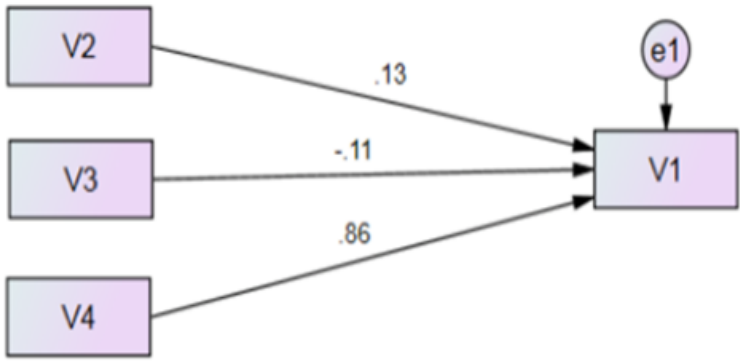

FIGURE 3. Multiple regression

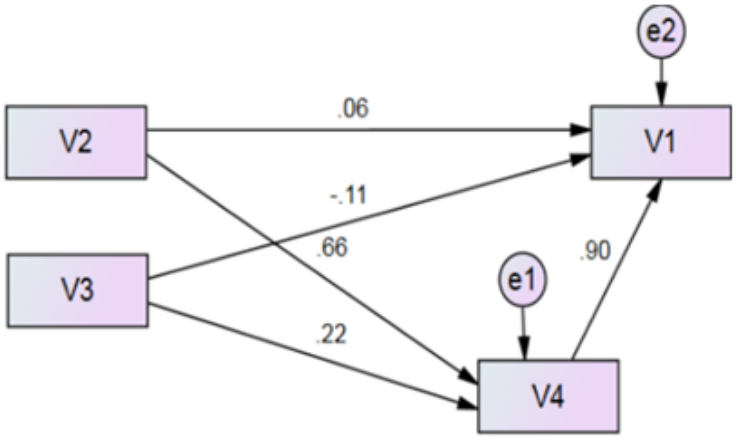

FIGURE 4. Path analysis model

	Estimate	Std. Error	t-value	Pr(¿—t—)
Intercept	-3.541e+03	1.963e+03	-1.804	0.0719
education	7.133e+02	1.710e+02	4.171	3.62e-05
Previous Experience	-1.671e+01	3.656e+00	-4.751	6.20e-06
Beginning Salary	1.756e+00	6.071e-02	28.932	¡2e-16

TABLE 1. coefficient of regression model

	Path analysis	Neural network
Experiment RMSEA	0.603	0.557
Test RMSEA	0.652	0.559

TABLE 2. RMSEA

4. CONCLUSION

In order to effectively tackle the growing complexity of researching, studying, modeling, and solving new problems in physics, engineering, medicine, biology, and many other fields, we have to develop new computational methods that are more closely related to human way of thinking. Now in this study, we have tried to look at updated methods and use them instead of the old ones to get results that are more accurate. Specifically, regression is an old way of examining the relationship between independent and dependent variables that only measures the relationship between independent and dependent variables. However, in path analysis, which is a better method than regression, in addition to the relationship between dependent and independent variables, the relationship between independent variables is also measured. Another advantage of path analysis over regression is that it shows the correlation between variables as direct effect, indirect effect, and total effect [1]. In addition, when the number of variables and their interaction with each other is large and complex, it can be said that the neural network has a high capacity and can learn more by increasing the number of inputs / outputs of the network, so that it works better than regression in some situation. However, due to the limitations mentioned in the path analysis method, at the beginning of this study, we compared it with neural network. Based on the analysis performed on the data under study, we concluded that the neural network method is preferred over path analysis because of pattern recognition, learning, generalization and data summarization. The results of this work may extend in many scenarios. One may use several mediator variable and compare the neural networks with the path analysis. It is also suggested that further research by combining the two methods of path analysis and neural network to create nodes and layers among the independent and mediating variables of the path analysis model which is called the hierarchical neural network and will yield more valid and accurate results [3].

References

[1] A. Sheikhi, *Analysis and statistical modeling*, Shahih Bahonar university press, (2019), (in persian).
[2] M. Hosseini, *Introduction to neural network*, (2007).
[3] M. B. Menhaj, *Computational intelligence*, Amir Kabir university, (2018).
[4] H. Breck, *Neural network computers finding practical application at lockheed*. J. Algebra, **91** (1990).
[5] J. Hopfield, *Neural networks and physical systems with emergent collective computational abilities*, Proceeding of the National Academy of Sciences of the USA, 2554-2558, (1982).
[6] R. Brian, *Pattern recognition and neural network*, Cambridge university press, **88**, (1996).
[7] R. Frank, *The Perceptron-a perceiving and recognizing automaton*, (1975).
[8] T. A. A, *Decision support system and intelligent system*, Prentice-Hall, **334**, (1998).

Fereshteh Aard
Department of Statistics, Shahid Bahonar University Kerman, Iran
Email: fereshteh_arad@yahoo.com

Ayyub Sheikhi
Department of Statistics, Shahid Bahonar University Kerman, Iran
Email: sheikhy.a@uk.ac.ir