Building a Predictive Model for Tobacco Production in Syria Using the Multi-Layer Neural Network (Perceptron MLP)

Majd Namaa[1], Ghassan Yacoub[2], Ayman Achouch[3]

[1]*Postgraduate student (MA), Dep. Agricultural Economics., Fac. Agric., Tishreen Univ., Lattakia, Syria*
[2]*Professor, Dep. Agriculture Economics., Fac. Agri, Tishreen Univ., Lattakia, Syria.*
[3]*Professor, Dep. Statistics and Programming , Fac. Economic., Tishreen Univ., Lattakia, Syria.*

Abstract: *The research aimed to build a predictive model for tobacco production at the level of the Syrian Arab Republic using neural networks technology, based on the statistics of the World Food and Agriculture Organization (FAO) for the period between (1975-2017), where the relationship between independent variables (year of production, area planted with tobacco was studied). Yield per hectare of tobacco crop) and its effect on the amount of production of the crop as a dependent variable based on the descriptive approach and the analytical statistical approach and building a multilayer neural network model (Perceptron MLP) and training it with a back propagation algorithm using statistical software (SPSS20).*

The neural network consisted of three layers (the input layer, the processing layer or the hidden layer, in addition to the output layer), where the input layer consisted of three units (the number of independent variables), while the output and hidden layers consisted of one unit. The input layer received the information from the time series data, the hidden layer processed the information and produced the output layer. The values from the input layer entering the hidden node are multiplied by weights, which is the set of pre-determined numbers and products are then added to produce a single number. She passed this number as a parameter to a nonlinear mathematical function (activation function) using the Hyperbolic function as an activation function in the hidden layer, and the Identity function was used in the output layer.

34-year data was used in the training of the chosen neural network, and 9 years in the process of network testing and comparison of network outputs with actual outputs.

The model showed high accuracy in prediction, as the total of error squares (0.18) in the training phase with a accuracy of (99.82)% and in the testing phase (0.17) with an accuracy of (99.83)%.

The neural network showed the relative importance of the independent variables in the impact on agricultural production of the tobacco crop, where yield per hectare occupied the first rank in the influence according to the predictive model by (100)%, while the cultivated area affected production by (77.2%), while the productive year was its effect Little by (4.7)%.

Key words: *Production, Tobacco, Syria, Multi-layer neural network*

I. INTRODUCTION

Tobacco, in its various varieties, is a crop of high economic value, and its cultivation has made remarkable progress in the country's agriculture. Its cultivation has spread since the beginning of the last century and was mainly concentrated in the coastal region. In addition, the state has given it special interest; as the establishment of the General Tobacco Corporation, which oversees the cultivation and marketing of this crop from different regions, according to predetermined plans and programs.

Tobacco is one of the most important crops as it contributes to pushing economic development through capital use, it is important to point out here that the value of the annual tobacco income in the Syrian national economy in (2016) amounted to (1904.8) million Syrian pounds [1].

It contributes to improving the quality of life by employing labor and reducing unemployment As the number of workers in the tobacco industry in Syria reached (11800) people in (2016) and more than 37,000 families in the same year worked in its cultivation. Therefore, it is of high economic importance compared to the main crops in the same region. Moreover, it is grown mainly in relatively poor geographical regions in its natural resources, as is the case in the mountainous coastal highlands, where the variety of Shk Al-Bent (Al-Baladi) is mainly cultivated. The most cultivated varieties in medium-altitude regions are Berlip, Basma, Granada, and Zegrin. While the Burley, Virginia, and Tanbak varieties are primarily cultivated in coastal areas [1].

Predicting the future values of time series has many important applications in many fields such as medicine, finance, engineering, market, weather, statistics, etc. One of these applications is to prevent an unwanted event from happening, by predicting it before it happens (such as forecasting a product that we want to sell on the market) and in another application predicting an unwanted event that is impossible to avoid, but its impact (such as storms) can be reduced. Forming a time series is usually the first step of a prediction that will provide an estimate of future values based on past values of the time series [2].

The Artificial Neural Networks (MLTP) have shown great success in the field of time series forecasting, hence the need to use and rely on them in particular due to their flexibility and reliability in time series forecasting [3].

The term "Neural network" is derived from the human brain, which contains parallel connections that achieve huge tasks in a short time compared to the computer, as the human brain contains highly complex nonlinear computational elements called neurons [4].

One of the fascinating paradoxes is that a computer's speed is 10 billion times higher than a neuron's speed, yet the average human can recognize a familiar face in a tenth of a second, faster than a computer that would require hours to recognize the face, so how can slow cells, compared to the computer, reach solutions at high speed?

Scientists' convincing and logical explanation is that the secret of the power of these networks lies in the way they process data; it processes its data in parallel, giving it a super-fast speed [4].

It turns out that the human neuron can be simplified and then mathematically represented to obtain artificial neurons.

Neuron networks consist of a group of nodes that perform a special kind of calculation collectively and that each node is a small standard arithmetic unit and that these nodes can work in parallel as they depend on interactions between them or how they relate and some scientists have identified it as:

Mathematical models that mimic the characteristics of biological systems that process information in a parallel manner, consisting of relatively simple elements called neurocytes [5].

It is a simple class or entity of mathematical algorithms that are formulated as graphs, these graphs are categorized into a large number of algorithms and these algorithms provide solutions to a number of complex problems.

The most prominent task of neural networks is the process of classification and coding, and the most prominent characteristics of neural networks are:

1- Noise resistance.

2- Its flexibility in dealing with distorted image systems.

3- Its ultimate resistance to distinguishing fragmented or partially degraded images [6].

The research was done because the following problem: The use of traditional forecasting methods results in high error rates, so using artificial neuron networks in this field gives high accuracy in predicting highly complex variables such as agricultural production as compared to traditional methods such as regression, etc.

Time series models are also frequently used to predict a variable's values; if the determinants of the variable are unknown, nor the factors that affect it, especially since agricultural production is subject to the influence of many factors that we cannot know its effect, such as natural factors, etc.

The importance of this research can be viewed from two aspects:

1- Economic aspect: Building a model for predicting tobacco production brings many economic benefits to the country; helping to draw up production-related plans such as cultivated areas, agricultural techniques, the number of workers and machinery, etc.

2- Statistical aspect: Statistical analysis based on modern models for time series prediction such as artificial neuron network models makes the model used more accurate compared to traditional models.

Based on the above, this research aims to:

1- Building a mathematical model to predict the variable of tobacco production based on time series data (1975-2017) using artificial neuron network.

2- Investigating the importance of artificial neuron networks in forecasting for the agricultural economic policymaker by increasing the degree of certainty and decreasing the degree of uncertainty by reducing the error value.

II. MATERIALS AND METHODS

A. Study place and data sources:

Tobacco production data for Syria for the period (1975-2017) were collected according to The Food and Agriculture Organization of the United Nations (FAO) statistics [7].

B. Research Methodology:

Descriptive and analytical method were used.

First: Descriptive statistical approach:

This method was used to describe time series data by presenting and summarizing it in statistical methods such as tables, graphs, and certain descriptive measures, such as measures of central tendency and measures of dispersion, to know the details related to that data, such as patterns of increase or decrease or persistence in the data, and to know the nature of the changes affecting the series.

Second: The analytical statistical approach:

Statistical methods were used to construct the model of artificial neuron networks, using some statistical programs such as SPSS20; Statistica; Excel.

A multilayer neural network (MTP) has been used due to its relevance to the nature of the data collected; This network consists of the input layer and one or two processing layer, so that the process layer does not exceed two layers. In addition to that, weights are adjusted to only one of the interconnecting layers that connect the previous layers so that the other layer (if any) remains constant in weights. The idea behind this network or how to teach these neuron networks is in two stages [8]:

1- Training phase

I Testing phase

Training phase:

It is the stage where the interconnect weights are adjusted to reach weights that are able to give correct answers, and this is done by the processing units doing three main operations:

a) Weighted Sum:

Each processing unit performs the addition of each input weight associated with the link that connects it with the unit in the layer that precedes it, multiplied by the value outside of that unit, which is as follows:

Sj =Ai wji

Where **wji** is the weight from node **i** (on the previous layer) to node **j** (on the current layer).

Ai is the value outside of unit **i**

Sj is the summation result for each unit of treatment **j**.

b) Transformation process:

This process takes place in the last layer of processing layers where the result of the addition process mentioned in the previous process is converted to one of the values that are supposed to be within the desired network products. For example, if the network were to learn how to classify numbers into odd and even, with each odd given the value 0 and each even number the value of 1.

The value of Sj, which is the product of the addition, will not usually give the value 0 or 1, so this output must be converted to one of these two values, via the conversion rule that the programmer defines. For example, the rule is as follows:

ifSj> 0 then Xj = 1

ifSj<= 0 then Xj = 0

Where Xj is the value output from the processing unit j [9]

c) Weights adjustment:

After completing the conversion process, the result that the grid gives is compared to the correct result that the network is supposed to give, by way of subtracting the target (correct) from the network output. The network needs to control its weights, through the following learning rule:

wjinew = wjiold + C (tj-Xj) ai

Where wjinew is the value of the new weight attached to the interface between unit j and unit i

Wjiold is the value of the old weight attached to the interface between unit j and unit i

C is the learning rate, which is a constant value, usually a value less than 1

Tj is the target value of the network

Xj is the value produced by the grid

And ai is the output from unit i

C. Training methods in neural networks:

Two methods can be used to train neural networks as a means of learning or simulation between the system and the user, and these methods are:

a- Supervised (training with a supervisor).

b-Unsupervised (training with no supervisor).

a. Supervised method:

It is assumed that a supervisor is present during the training process for each model and is used for testing and takes place during (target output) and includes the input model in addition to the desired output to determine (real output). In this process, a comparison is made between the real output until the performance of the network is reduced after giving a matrix of weights and determining a vector of error (the real vector), where it compares the output vector with the output and the input vector of the network, and also compares with the desired output to identify errors through the following equation:

Real output- target output = error

a) Unsupervised method:

In this method, there is no supervisor to provide the required models, so the system must learn by exploring and how the properties or structured factors in the input model for the experimental user and this learning must be done through strengthening the weights selected for the nodes in order to match the experimental models and distinguish learning in this way in a repeated way until it is completed Weight stability.

D. Neural network architecture:

-single layer network

This network has a single layer of weights, either the input units can be defined as the signals received from the outside world. These units are linked to the output units (which represent the answer to the input units) and are fully correlated, for example, an auto hetor Adeline preceptor with a backward flow character, and the output units are not linked to any other units except Hopfield network in which all the input units are the same as the output units, Figure (1) represents an example of these networks [Hum92] [Kit90].

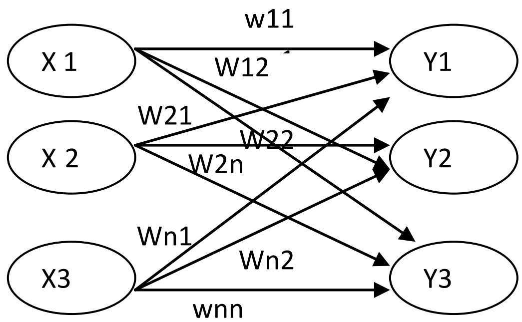

Figure (1): The architecture of single layer network

-Multi layer network

These are networks that have one or more layers of hidden nodes that can be defined as units that do not represent input units or output units. This network can solve more complex problems of the first type (single one), but its training is more difficult.. Medline, Back propagation necogmatron miffs is one of its examples. Figure (2) represents the model of these networks.

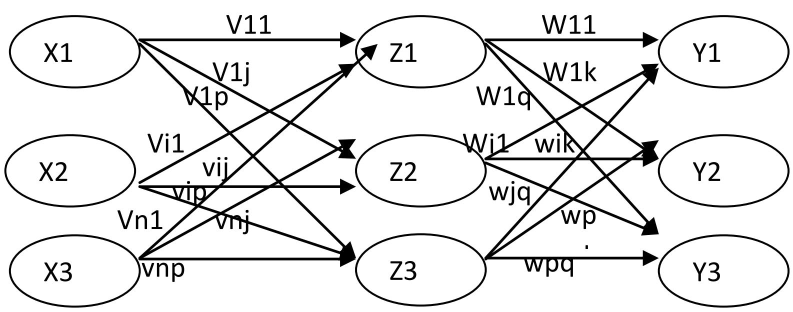

Figure (2): The architecture of the layered network

E. Study Variables:

- The Independent Variables:

Year of production: a time series was taken (1975-2017).

The cultivated area: where the data of the cultivated area of tobacco crop were used at the level of Syria, based on (FAO) statistics during the period (1975-2017) [3].

Yield: where the data on yield per hectare of tobacco crop were used, depending on the (FAO) statistics during the studied period [7].

- The Dependent Variable:

Tobacco production: The quantity of agricultural production from tobacco crop during the period (1975-2017) was used, and the neural network model was built to predict agricultural production by taking the relationship between the dependent variable and the independent variables.

So the Research hypothesis was to predicting the future values of agricultural production using neural networks is among the best methods used for predicting the future, such as regression and others.

III. RESULTS AND DISCUSSION:

A. Descriptive analysis of the study variables

Table (1). Statistical description of research variables during the period (1975-2017)

the variable	the average	standard deviation	Coefficient of variation%	the biggest value	the least value	range
the production (ton)	17621.20	5472.820	31.05	28870	10912	17958
the area (h)	13770.14	2355.419	17.10	18100	7890	10210
the yield (ton/h)	1.29	0.31	24.03	1.77	0.68	1.09

Source: Calculated by the researcher based on FAO statistics (1975-2017) using SPSS program

Results in Table (1) showed the statistical description of the economic variables under study, as production reached its lowest level in (2017) by about (17958) tons, and the highest level in (2005) by about (28870) tons, with an average of 17621.20 and a standard deviation of 5472,820. The value of the coefficient of variation is 31.05%, meaning that the difference between the studied data is 31.05%, and the homogeneity is 68.95%

While the cultivated area ranged between 7890 hectares in (2017) and 18100 hectares in (2000), with an average of 13770.14 tons and a standard deviation of 2355.419, and the value of the coefficient of variation was 17.10%, meaning that the percentage of difference between the studied data is 17.10%, and homogeneity is 82.90%.

The yield per hectare ranged between 0.68 (tons / ha) in (1976) and 1.77 (tons / ha) in (2005), with an average of 1.29 and a standard deviation of 0.31, and the value of the coefficient of variation was 24.03%, meaning that the percentage of difference between the studied data is 24.03% , and homogeneity of 75.97%.

Table (2): Matrix of correlation between studied variables:

		time	production	cultivated area	yield
time	Pearson Correlation	1	.405**	-.471**	.806**
	Sig. (2-tailed)		.007	.001	.000
production	Pearson Correlation	.405**	1	.497**	.807**
	Sig. (2-tailed)	.007		.001	.000
cultivated area	Pearson Correlation	-.471**	.497**	1	-.101
	Sig. (2-tailed)	.001	.001		.520
yield	Pearson Correlation	.806**	.807**	-.101	1
	Sig. (2-tailed)	.000	.000	.520	

Source: Calculated by the researcher based on FAO statistics (1975-2017) using SPSS program

Results in Table (2) demonstrated the following:

1- There is an medium positive significant correlation between time and production, where the value of the correlation coefficient between the two variables was (0.40) with a statistically significant level of significance (0.007 <0.05).

2- There is an medium negative significant correlation between time and the cultivated area, where the value of the correlation coefficient between the two variables reached (-0.47) with a statistically significant level of significance (0.001 <0.05).

3- There is a large positive significant correlation between time and yield, where the value of the correlation coefficient between the two variables was (0.80) with a statistically significant level of significance (0.000 <0.05).

4- There is an medium positive significant correlation between production and the cultivated area, where the value of the correlation coefficient between the two variables was (0.49) with a statistically significant level of significance (0.001 <0.05).

5- There is a large positive significant correlation between production and yield, as the value of the correlation coefficient between the two variables reached (0.80) with a statistically significant level of significance (0.001 <0.05).

6- There is a low negative correlation between the cultivated area and yield, as the value of the correlation coefficient between the two variables reached (-0.10) with a significant level not statistically significant (0.52 <0.05).

B. Designing the neural network:

The multi-layer module (Perceptron MLP) was used using the statistical software (IBM SPSS 20) to build the neural network model and test its accuracy. 34 years data were used in the training phase by (79.1)%, and 9 years for the testing phase by (20.9%) (Table 3).

Table (3): Summary of the processing procedure in a neural network:

Case Processing Summary			
		N	Percent
Sample	Training	34	79.1%
	Testing	9	20.9%
Valid		43	100.0%

Source: the output of the neural network using SPSS

Results in Table (4) shows information about the neural network. The number of units in the input layer is the number of independent variables (year of production, cultivated area, yield) in addition to the total number of factor levels. A separate unit is created for each class of education and none of the classes are considered "redundant" units as is usual in many modeling procedures. Likewise, one unit is selected in the hidden layer and so is one unit in the output layer.

The Hyperbolic function (hyperbolic shadow) was used as an activation function in the hidden layer which was gives by the following relation:

$$\gamma(c) = \tanh(c) = \frac{e^{c} - e^{-c}}{e^{c} + e^{-c}} + Bias$$

where: y: the dependent variable, c: the independent variable.

Also, The Identity function (the neutral or identical function) is used in the output layer. It is a function in which each element is related to itself, or which the domain and the corresponding domain are the same group. This is given by the following relationship:

$$y(c) = c + Bias$$

Table (4): Information about a multi-layered neural network:

Network Information			
Input Layer	Covariates	1	T
		2	Area
		3	Yield
	Number of Units[a]		3
	Rescaling Method for Covariates		Standardized
Hidden Layer(s)	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1[a]		1
	Activation Function		Hyperbolic tangent
Output Layer	Dependent Variables	1	Product
	Number of Units		1
	Rescaling Method for Scale Dependents		Standardized
	Activation Function		Identity
	Error Function		Sum of Squares
a. Excluding the bias unit			

Source: the output of the neural network using SPSS

Figure (3) shows the architecture of a neural network, which is a three-layer network of interconnected nodes: an input layer, a hidden layer, and an output layer. The nodes between the input and output layers can form one or more hidden layers. Each neuron in one layer has a connection to every other neuron in the next layer, but the neurons belonging to the same layer have no connections between them. The input layer receives information from time series data, then the hidden layer processes information, then produces the output layer and predicts continuous values. Values from the input layer which entering the hidden node are multiplied by weights, and these are the predefined numbers, then products are added to produce one number. This number is passed as an parameter to a nonlinear mathematical function (the activation function) where the Hyperbolic function was used as an activation function in the hidden layer [10], and the Identity function was used in the output layer as shown in the figure (3).

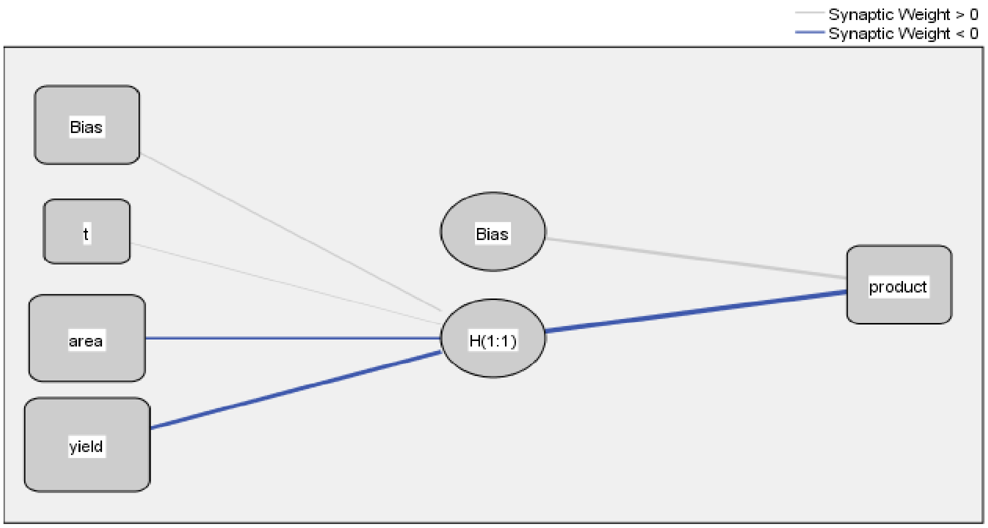

Figure (3): The architecture of the multi- layer neural network (Perceptron MLP)

Table (5): Summary of the neural network model used

Model Summary		
Training	Sum of Squares Error	.186
	Relative Error	.011
	Stopping Rule Used	1 consecutive step(s) with no decrease in error[a]
	Training Time	0:00:00.00
Testing	Sum of Squares Error	.178
	Relative Error	.024
Dependent Variable: product		
a. Error computations are based on the testing sample.		

Source: the output of the neural network using SPSS

Table (5) shows information related to the training and testing results. A cross-entropy error is given for both the training and the sample test since the error function that the network reduces is during the training phase. As the sum of squares of error in the training phase was (0.186). This error indicates the ability of the model to predict production. The entropy error is lower for the test sample compared to the training data set (0.178), and this means that the network model was not well equipped with the training phase.

Table (6): Parameter estimates used in the training and testing phases

Parameter Estimates			
Predictor		Predicted	
		Hidden Layer 1	Output Layer
		H(1:1)	Product
Input Layer	(Bias)	.331	
	t	.030	
	area	-.351	
	yield	-.560	
Hidden Layer 1	(Bias)		.491
	H(1:1)		-1.884

Source: the output of the neural network using SPSS

Results in Table (6) showed the values of the parameters used in building the predictive model in the input layer, where three variables were used (year of production, cultivated area, yield) in addition to the bias parameter, and the value of this parameter also appears in the hidden layer. The idea of bias comes that neural networks are inspired by biological neurons. The basic idea is that human neurons take a bunch of inputs and add together. If the sum of the inputs is greater than some thresholds, the neuron will produce an output that goes to the other neuron cells. This threshold is basically the same as the bias. In this way, bias in artificial neural networks helps replicate the behavior of real human neurons [11].

Depending on table (6), we can deduce the activation function in the hidden layer:

$$\gamma(c) = \tanh(c) = \frac{e^{0.30t-0.351a-0.560yi} - e^{-(0.30t-0.351a-0.560yi)}}{e^{0.30t-0.351a-0.560yi} + e^{-(0.30t-0.351a-0.560yi)}} + 0.331$$

As well as the matching function in the output layer:

$$y = -1.884c + 0.419$$

where: t: year of production, a: the cultivated area, yi: yield.

Table (7): the relative importance of the independent variables

Independent Variable Importance		
	Importance	Normalized Importance
t	.026	4.7%
area	.424	77.2%
yield	.550	100.0%

Source: the output of the neural network using SPSS

Table (7) explains the relative importance of the independent variables in the model, where yield took the highest percentage in importance (100%), the cultivated area was (77.2%), while the effect of the production year was small at (4.7%). Figure (4) confirms the validity of the results obtained.

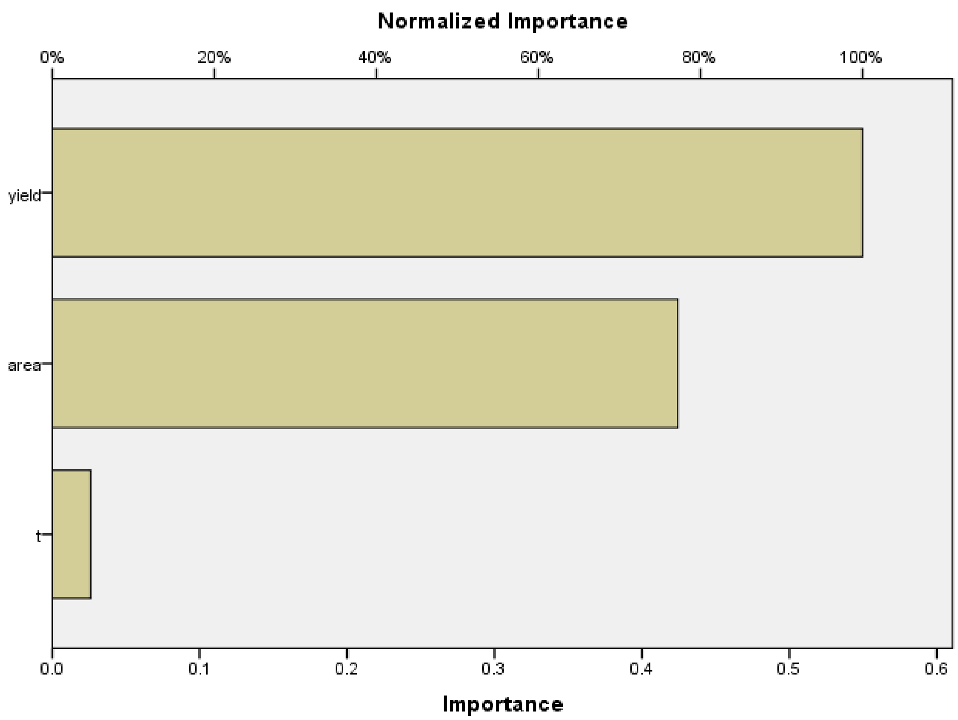

Figure (4): the relative importance of the independent variables.

IV. CONCLUSION

-The neural network technology is the most accurate statistical method, as it can deal with any type of data and thus discover a suitable model for the data while minimizing the error.

-The neural networks method deals with many different cases, so the degree of its complexity is more than it is with other prediction methods.

-The neural networks method requires the use of the computer for a considerable period of time in the language of computers, which makes it very expensive on the other hand, as this method may be more accurate than other prediction methods, and here we face the choice between accuracy and cost.

-The neural network showed the relative importance of the independent variables in affecting agricultural production from tobacco crop, as the yield of one hectare occupied the first place in the effect according to the predictive model at a rate of (100)%, while the cultivated area affected the production by (77.2)%, and the productive year was its effect a small percentage (4.7%).

REFERENCES

[1] Tobacco balance sheet, 2017. The Ministry of Industry. Official Statistics Office. Syria

[2] John Wiley, Sons (2008). "*Introduction to Time Series Analysis"*

[3] Ali Abdul-Azim Abdul-Karim (2013), "*Using Artificial Neural Networks to Predict a Multidimensional Macroeconomic Model in Iraq during the Period (1996-2007)"* (Gulf Economic Journal, Issue (24).

[4] Ali AwadallahTayfour (2010), "*Introduction to Artificial Neural Networks*" (College of Engineering, Sudan University of Science and Technology).

[5] Yochanan Shachnurove , " *Utilizing Artificial Network to predict Stock Markets* ", the city collage of the city University of New York & The University of pennsylvania , September 2000 p.1

[6] Leonldas Anastasakis , &Neil Mort, "*Network-based of the USD / GBP exchange rate the Utilization of data compressionTechniques For input imension Reduction* " , University of Sheffield , SI 3JD , U.K , p.1

[7] www.FAO STATISCS.com

[8] Al-Kaabi Haider Hamid, Ali Darb Al-Hayali (2015), "*Predicting the Food Gap for Wheat and Rice Crops in Iraq Using the Artificial Neural Networks Method ANN for the period 2014-2022*", (Al-Muthanna Journal of Agricultural Sciences, Iraq, Volume 4, page 1-21

[9] Coryn AL-Baler-Jones Daviad J.C Mackag " *A Recurrent Neural Network for Modeling Dynamical* ", Department of Materials Science &Metallurg , University of Cambridge, Pembroke S.T,Cambridge ,CB23 QZ,England , 1998,p.Forecasting ", Technical University of Denmark.

[10] K.WAR Wick ; QMZHU & ZMA ," *A hyperstable neural network for the modeling and control of nonlinear system* " . Sadhand ,vol.25 , part 2 , April 2000 , p169 – 180 , printed in India

[11] Khaled Abdel-Wahid Mohamed Ahmed (2001), "*Application of artificial neuronal networks to predict the rate of self-sufficiency of red meat"* (Journal of Agricultural Sciences, Faculty of Agriculture, Mansoura University, Egypt)

[12] Gurpreet Kaur, Gurmeet Kaur" *Various Types of Functional Link Artificial Neural Network Based Nonlinear Equalizers used in CO-OFDM System for Nonlinearities Mitigation*" SSRG International Journal of Electronics and Communication Engineering(SSRG - IJECE)–Volume5 Issue2 Feb 2018.