TÜRK TARIM ve DOĞA BİLİMLERİ DERGİSİ

TTDB

TURKISH JOURNAL of AGRICULTURAL and NATURAL SCIENCES

www.dergipark.gov.tr/turkjans

Research Article

Application of Artificial Intelligence Methods to Predict Cotton Production in Turkey

Nazire MİKAİL[1] Mehmet Fırat BARAN[2*]

[1]Department of Animal Science, Faculty of Agriculture, Siirt University, 56100 Siirt-TURKEY
[2*]Department of Biosystem Engineering, Faculty of Agriculture, Siirt University, 56100 Siirt-TURKEY
Corresponding Author: firat.baran@siirt.edu.tr

Abstract

Farmers are always curious about the factors affecting yield in plant production. Determining these factors can give information about the yield in the future. Reliability of information is dependent on a good prediction model. According to the operating process artificial neural networks imitate the neural network in humans. The ability to make predictions for the current situation by combining the information people have gained from different experiences is designed in artificial neural networks. Therefore, in complex problems, it gives better results than conventional statistical methods. In this study, artificial neural networks and support vector machines methods of artificial intelligence were used in order to predict the production of cotton. From a comprehensive data collection spanning 73 farms in Diyarbakır, Turkey, the mean cotton production was prevised at 559.19 kg da-1. There is four factors that picked as pivotal input into this model. As a result, the ultimate artificial neural network model is able to foreshow cotton production, which is built on elements like: farm states (cotton area and irrigation periodicity), machinery usage and fertilizer consumption. At the end of the study, cotton yield was estimated with %84 accuracy.

Anahtar kelimeler: Crop fields, yield, cotton, ANN, SVM, Turkey

Türkiye'de Pamuk Üretiminin Tahmini İçin Yapay Sinir Ağı Yöntemleri Uygulaması

Öz

Yetiştiriciler bitkisel üretimde verimi etkileyen faktörleri hep merak etmişlerdir. Bu faktörlerin belirlenmesi ileriedeki verim hakkında bilgi verebilir. Bilginin güvenilirliği iyi bir tahmin modeline bağlıdır. Çalışma sürecine göre yapay sinir ağları, insandaki sinir ağını taklit eder. Yapay sinir ağlarında insanların farklı deneyimlerden edindiği bilgileri birleştirerek mevcut duruma yönelik tahminler yapabilme yeteneği tasarlanmıştır. Bu nedenle karmaşık problemlerde yapay sinir ağlarına göre daha iyi sonuç verir. Bu çalışmada, pamuk üretimini modellemek için yapay sinir ağı yöntemi kullanılmıştır. Türkiye, Diyarbakır'da 73 işletmeyi kapsayan kapsamlı bir veri koleksiyonundan, ortalama pamuk üretimi 559,19 kg/da olarak hesaplanmıştır. Bu modele temel girdi olarak seçilen dört faktör vardır. Sonuç olarak, nihai YSA modeli, işletme durumları (pamuk alanı ve sulama periyodu), makine kullanımı ve gübre tüketimi gibi unsurlara dayanan pamuk üretimini gösterebilmektedir. Çalışma sonunda pamuk verimi %84 doğrulukla tahmin edilmiştir.

Key words: Tarla ürünü, verim, pamuk, ANN, SVM, Türkiye

Introduction

Cotton is one of the most prominent products in the agriculture sphere, industry and trade because of its very different and important utilization areas. In addition to the escalating world population, the increasing needs of human beings for consumption raise the importance of this versatile plant day by day. Growing interest in natural fibers and rising living standards in the world increases the demand for cotton plants (Anonym 2020 a).

According to the data of the International Cotton Advisory Committee (ICAC), 32,825 million hectares of cotton were produced in the world in the 2018/19 production period. In this season, 37% of the 32,825 million hectares of cotton cultivated in India. India is followed by the USA, China, Pakistan and Brazil in the width of the cultivation areas.As the result of the expansion of cotton acreage in African countries in recent years, despite the grow, Turkey has been ranked 11th in terms of world cotton cultivation area (Anonym 2020b).

It is an important industrial plant that constitutes the raw materials with fiber in textile industry, oil obtained from its core in vegetable oil industry, achaenium and pulp in animal feed industry, linters in paper, furniture and cellulose industry. Cotton is an important and strategic product that provides great benefits to our country's economy with this wide area of use, added value and employment opportunities. Due to these features, it has a contribution to the development of both agriculture and industry of the regions and countries grown (Anonym 2020c).

Areas where have intensive cotton production in Turkey; Aegean, Çukurova, Southeastern Anatolia Regions and Antalya. In the 2017/18 cotton season, in 502 thousand hectares 882 thousand tons of cotton fiber production has been made, and about 1 million 571 thousand tons of cotton has consumed in Turkey. In the 2017/18 cotton production season, in return to 882 thousand tons of fiber cotton, 2,5 million tons of seed cotton was produced, and fiber cotton yield was 1820 kg / ha. Şanlıurfa, Aydın, Bursa, Diyarbakır, Adana and Izmir are 6 provinces in Turkey that meets 88% of production, respectively. Şanlıurfa province alone meets 42% of all production. The share of the other 23 cotton producing provinces in production is between 0.1% and 1.3% (Anonym 2020 d).

Since cotton is a selective plant in terms of climate characteristics, it can be grown in limited places in our country (Karademir et al. 2015). Cotton plant has great economic importance for humanity with its widespread and compulsory use, and for producer countries with the added value and employment opportunities that it creates. Increasing population, growing interest in natural fibers and rising living standards raise the demand for cotton plants.

Artificial neural networks have been used with traditional statistical methods in recent years thanks to its highly accurate prediction and classification capability and now it is even more popular than statistical methods. However, ANN modeling also has its own challenges. We can practice many potential ways to form and train networks. To put it in different way, with a sufficiently great number of independent parameters, neural networks is able to be trained to tightly suit data that is likely include noise. Hence, it is very important for us to understand how to optimise the structure of networks and learning in order to develop reliable models for decision-making and that generalize well to data that do not fit any distribution (Samarasinghe, 2006, p. 5).

In recent years, ANN models have been applied to predict the production of different plants using artificial neural networks. As an example, a high-accuracy production estimation model has been developed for basil (Rostami et al., 2017) and wheat (Safa et al., 2015). In addition, there are many studies that created ANN models that.measure the effect of potato production on the environmental quality index (Khoshnevisan et al., 2013a), predict the maturity of cotton fibers (Farooq et al., 2018), and estimate the massiveness of energy ingested in production (Safa & Samarasinghe, 2011; Taki et al. , 2012; Khoshnevisan et al., 2013b; Nabavi-Pelesaraei et al., 2016; Khoshroo et al., 2018; Taki et al., 2018). With the same success, ANN is being used in other fields of agriculture. For example, high accuracy models can be created in the estimation of the temperature inside the greenhouses (Saltuk & Mikail, 2019), in the beef cattle production (Bozkurt et al., 2015), in the estimation of milk yield (Mikail et al., 2013; 2014; 2016) and in the diagnosis of diseases such as mastitis (N. M. Mammadova & Keskin, 2015).

In this study, we have a purpose to set down the elements affecting cotton production in Diyarbakır, to create a model with the help of these factors and to make a model with support vector machine and artificial neural networks. Also, it is aimed to make yield estimation..

Materials and Methods

Diyarbakır province is in the central part of the Southeastern Anatolia Region and at the northern end of Mesopotamia. It is surrounded by the provinces like Siirt, Muş from the east; Mardin from the south; Şanlıurfa, Adıyaman, Malatya from the west; Elazig and Bingöl from the north. Its area is 1.516.200,00 square kilometers, between 37.905199 and 40.231934 north latitudes and 40.37 and 41.20 east longitudes. It is surrounded by mountains that are not too high and in the middle, there is concavity. It is covered with 37% of mountains and 31% of lowlands. The lowlands are firtile and suitable for agriculture. These fecund lands are irrigated by the Tigris river and its

tributaries. The city was founded on a horizonal surface on the eastern edge of the broad basalt plateau stretching between Karacadağ and the Tigris, above the Tigris valley and at the top of the river curve. Its altitude is 650 meters above sea level. This altitude changes between 640 m and 660 m in some places (Anonym, 2020e)

The core material of this particular study is the data derived from cotton producing agricultural enterprises in Bismil district of Diyarbakır. The data covered by the research were obtained through 73 face-to-face surveys in 2019. However the secondary data of the research were used from the publications of various national and international institutions and organizations related to the subject.

Method

Sampling Method

The Simple Random Sampling Method was used to determine the sampling frame and sample number in the study (Yamane, 1967).

$$n=(N \times s^2 \times t^2)/(N-1)d^2+(s^2 \times t^2) \quad (1)$$

n = Sample Size
s = Standard Deviation
t = "t value" Related to the Selected Confidence Limit
N = Total Number of Units for Sampling Frame
d = Acceptable Margin of Error (%)

In the study, it was deemed appropriate to conduct a questionnaire on cotton production in 134 enterprises with 95% confidence interval and 5% deviation from the mean.

The Farmer Registration System of Diyarbakır Provincial Directorate of the Ministry of Agriculture and Forestry records were used as a sample villages and enterprises.

The primary data to be used in the study were the data obtained from the cotton producing enterprises in the sampling area. In determining the settlements within the scope of the survey application, the indicators of the ability to represent the research area were used, taking into account the characteristics such as cotton cultivation area, number of enterprises, agricultural production potential, presence of tools - equipment and etc. The questionnaire forms used were arranged in accordance with the characteristics of the research field. Before the survey application, trial questionnaires were conducted in the selected settlements, after that we have given the final to the questionnaire forms considering the deficiencies observed.

Withal, these oblique details and cotton production were surveyed to design the model in order to be able to prognosticate cotton production. To be able to implement this in the ANN model, it is essential to pick a finite number of relevant and significant variables, voiding any sort of propensity. For that reason particularly, all of the information was researched cautiously. 17 original variables were participated in order to be presented as a most likely input in the final model. Questionnaires containing unreliable information are not included in the process. For example, irrigation frequency entered as 0 is excluded. We used this information to draw the graphs and carry out statistical analysis operating with MINITAB Statistical Programme and MATLAB Software (Mathworks, 2009). In this context, 73 surveys conducted in Diyarbakır province Bismil central districts and villages were evaluated and estimation application was made.

Support Vector Machines

Support vector machines (SVM) proposed by Vapnik (1996, 1998), are characterized by an internal process of constructing classification rules quite different in basic concept from those of the statistical methods. SVMs are often effective in cases for which ordinary classification methods are not effective, such as problems based on high-dimensional data and data with nonlinear structure, and they have been adapted and applied in many fields (Konishi, 2014).

Suppose that we have n observed training data x1, x2, …, xn relating to p variables x=(x1, x2, …, xp)^T that characterize certain individuals or objects. Here, let us assume that we already know the membership of the training data in the two classes G1 and G2. Our objective is to find a hyperplane

$$w_1x_1+w_2 x_2+\cdots+w_p x_p+b=w^T x+b=0, \quad (2)$$

which separates two classes, where w=(w1, w2, …, wp)^T.
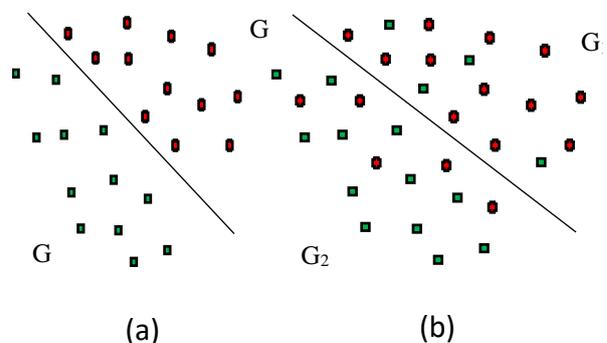

Figure 1. Linear separable data by a hyperplane (a), non-separable data by any such linear hyperplane (b)

The meaning of inherently obvious separability in the training data is illustrated by the configuration in Fig.1(a). As shown in the figure, the training data are completely separable into two classes by a hyperplane, which in this case is a line because the data space is two dimensional. This is known as linear separability. For the configuration in Fig. 1(b), separation into two classes cannot be obtained by any such linear hyperplane (Konishi, 2014).

Artificial Neural Network

In Fig. 2 ANN model with one input, one hidden neuron and one output was pictured respectively in input layer, hidden neuron layer and output layer. As a result, this ANN model has a_1 - one input-hidden layer weight and b_1 - one hidden-output layer weight. Here, X is the input and the z is the network output. Bias input of + 1 has noted in the hidden neuron and output neuron, respectively with an associated weight of a0 and weight of b_0. Smith also used this notation for the weights.
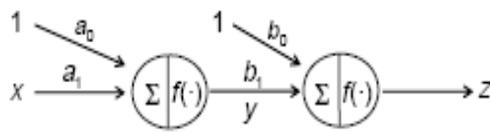

Figure 2. One-hidden neuron, one-input and one-output neural network

The network feeds with the input and the hidden neuron computes the weighted total of inputs (covering also bias) and delivers this weighted total of inputs via the logistic function to be able to built the hidden-neuron output, y. However, the output neuron feeds by the output neuron y as input via the associated connection link. By this link it is being weighted. After that, the weighted input is transferred by the activation function of neuron. Network output appears as a result of the output of this neuron's transformation. The hidden neuron supersedes very significant part of processing.

The weighted input is shown below:

$$u=a_0+a_1 x \qquad (3)$$

After the first step, it follows by the idea the weighted sum u should be passed through the logistic function by the hidden neuron. U is the logistic function's argument and this function, without fail, is a standard function with y=0.5 at u=0. It will be more beneficial for the output y to be specified in terms of input x, so, it will be possible to portray y is being laid out by x through u. Replacing u inside the logistic function, output y as the hidden-neuron is also shown below:

$$y=1/(1+e^{-(a0+a1 x)} \qquad (4)$$

Model Assessment

When more than one model are used to modeling the data, the fit of the models to the data is tested via the goodness of fit criteria. In this model, we used r- Pearson correlation coefficient, R2 - coefficient of determination, RMSE - root mean squared error, MAE - mean absolute error, MRAE - mean relative absolute error, MRSE - mean relative squared error, MSE - mean square error criteria for model comparison (Spiegel et al., 2009; Willmott & Matsuura, 2005).

a) The coefficient of correlation

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x}_i)(y_i - \bar{y}_i)}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x}_i)^2 \sum_{i=1}^{n}(y_i - \bar{y}_i)^2}} \qquad (5)$$

b) The coefficient of determination

$$R^2 = 1 - \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 / \sum_{i=1}^{n}(y_i - \bar{y})^2 \qquad (6)$$

c) Root mean squared error

$$RMSE = \sqrt{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 / n} \qquad (7)$$

d) Mean absolute error

$$MAE = \frac{\sum_{i=1}^{N}|y_i - \tilde{y}_i|}{n} \qquad (8)$$

e) Mean relative absolute error

$$MRAE = \frac{\sum_{i=1}^{N}\frac{|y_i - \tilde{y}_i|}{\tilde{y}_i}}{n} \qquad (9)$$

f) Mean relative squared error

$$MRSE = \frac{\sum_{i=1}^{N}\left[\frac{(y_i - \tilde{y}_i)}{\tilde{y}_i}\right]^2}{n} \qquad (10)$$

g) Mean square error

$$MSE = \frac{\sum_{i=1}^{N}(y_i - \tilde{y}_i)^2}{n} \qquad (11)$$

where, y_i – observed value, $\hat{y}_i$ – predicted value, $\bar{y}$ – the arithmetic mean, n – the total number of observations.

Results and Discussion

In the present study, the maximum and minimum yield varies between 410 and 680 kg per da, and the average yield was determined as 599.19 kg da-1 (Table 1). Average cultivation area in the enterprises was found to be 87.64 da. The

amount of fertilizer used in the enterprises was 87.84 kg per da, the tractor usage time was 1.8 hours and the irrigation frequency was approximately 6 times (Table 1).

Table 1. Descriptive statistics of variables used in the model

Variable	N	Mean	StDev	Min	Median	Max
Cotton area (da)	73	87.64	47.23	12.00	80	245
Fertilizer (kg da^{-1})	73	87.84	17.95	40.40	86.5	123.50
Tractor usage (h da^{-1})	73	1.80	0.15	1.47	1.79	2.12
Irrigation frequency	73	6.30	1.63	4.00	6.00	10.00
Yield (kg da^{-1})	73	559.19	70.60	410.00	570	680.00

In the study, no linear relationship was found between yield and cultivation area and tractor operating times. It was determined that the amount of fertilizer and the frequency of irrigation positively affect the yield, though it is weak. That is, the Pearson correlation coefficient between fertilizer amount and yield was $r = 0.29$ ($p < 0.05$), and the Pearson correlation coefficient between irrigation frequency and yield was calculated as $r = 0.38$ ($p < 0.01$) (Figure 3). Similar results were found in the study (Safa et al., 2015). In this study conducted in New Zealand, Pearson correlation coefficient between N fertilizer consumption and wheat yield was determined as $r = 0.43$
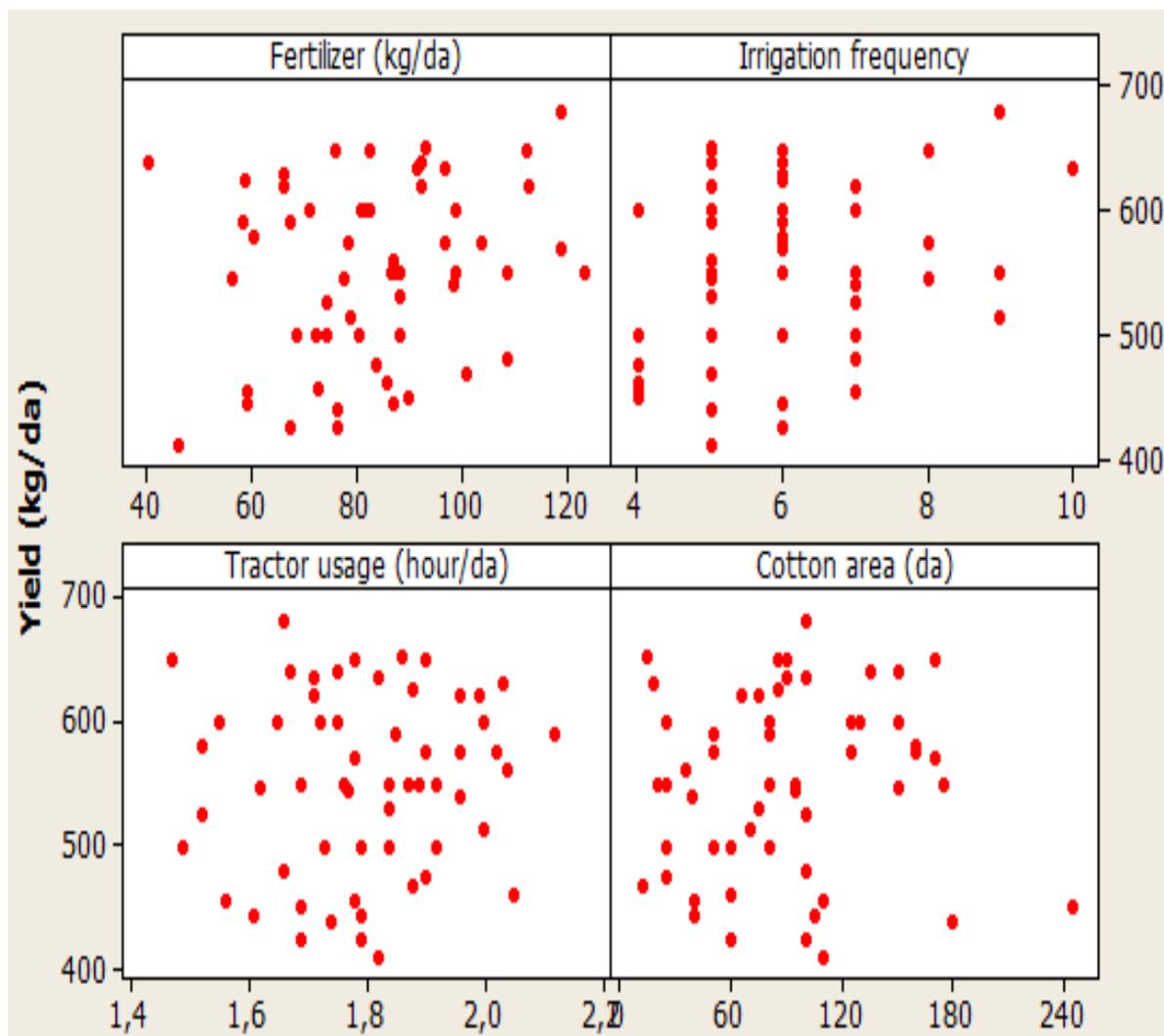

Figure 3. Correlations between cotton yield and input variables

Support Vector Machine Model

Various kernel functions can be applied in SVM modelling. RBF is the most popular kernel function used in regression problems (N. Mammadova & Keskin, 2013; Nguyen, 2019; Yingli et al., 2020). In order to estimate the cotton production, the RBF kernel function was used in SVM model. There were 49 support vectors created in the model.
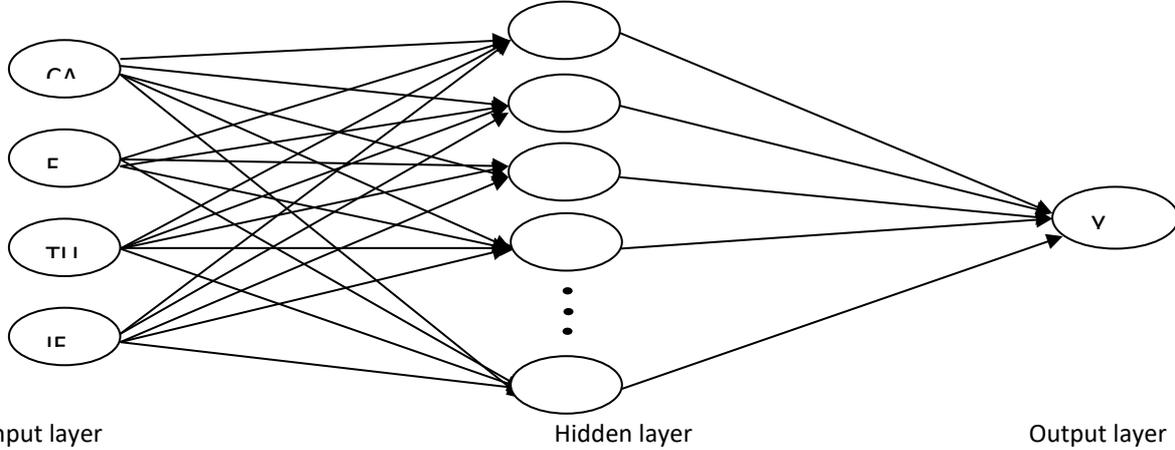
Input layer   Hidden layer   Output layer

Figure 4. Generated ANN model

Artificial Neural Network Model

According to this study, the ANN model consists of 20 neurons hidden layer, 4 neurons input layer and 1 output layer. (Figure. 4). Levenberg-Marquardt backpropagation learning method was used in the model. The data set is randomly divided into 3 shares: 70% training, 15% validation and 15% testing. The compatibility of the cotton production values estimated with each data set with the actual observation values was measured by the Pearson correlation coefficient (Figure. 5).
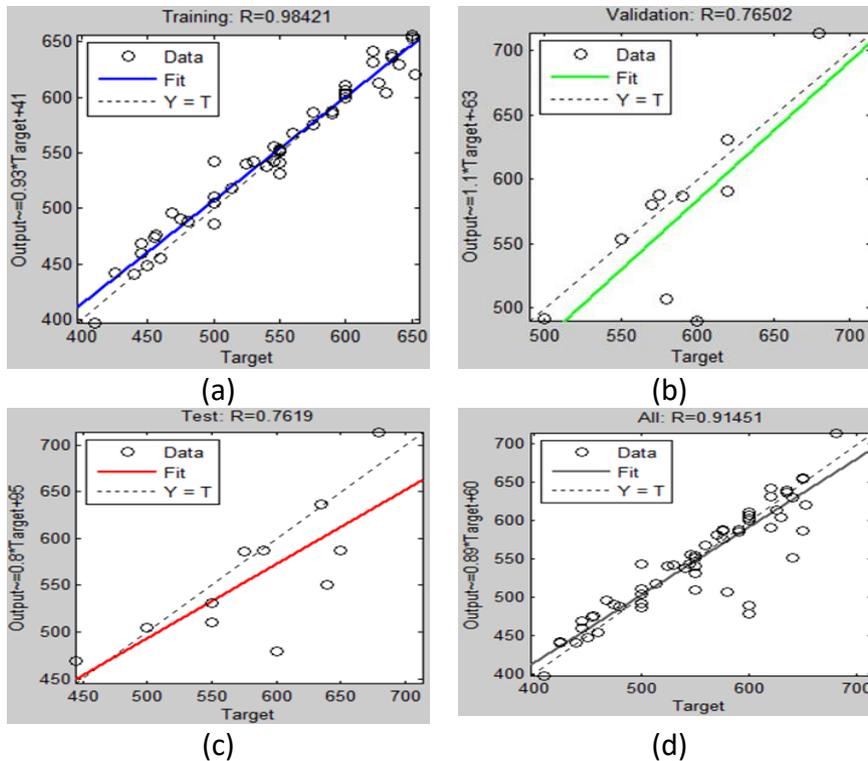

Figure 5. Correlation between observed and predicted cotton yield for a) training b) validation c) test and d) all data.

The model gave the best result after the 18th iteration. As can be seen from Figure 3, the model gave the best result with training data. Here, the degree of the linear relationship between the estimated yield values obtained by using the training data and the actual yield values, was found to be 0.98. This is followed by validation (0.77) and test (0.76) accordingly. When all data were evaluated together, this coefficient was found to be 0.91. Howbeit, the RMSE statistic of the model was found to be 20.37. This value was calculated as 44.81 in the estimates obtained by SVM (Table 2).

Table 2. Goodness of fit for models

	SVM	ANN
r	0.536	0.9145
R^2	0.29	0.84
MRAE	0.074	0.0322
MAE	41.909	17.2064
MRSE	0.011	0.0031
RMSE	44.81	20.37
MSE	3591.154	829.6069

As a result of the comparison between the ANN model and the SVM models, it was seen that the SVM model cedes the correlation coefficient between the exact and predicted cotton yield of the ANN model, as the ANN model was much higher. In addition, the RMSE value was found to be much lower than the value obtained from the SVM estimation (Table 2). As shown in Figure 6, the ANN model can predict cotton yield with an 84% accuracy. It was calculated as 91% in a similar study created for wheat production. The fact that there are weak relationships between the data leads to a very low accuracy as a result of the model made with SVM. In modeling with ANN, we can see that high accuracy predictions can be made even by using such data.
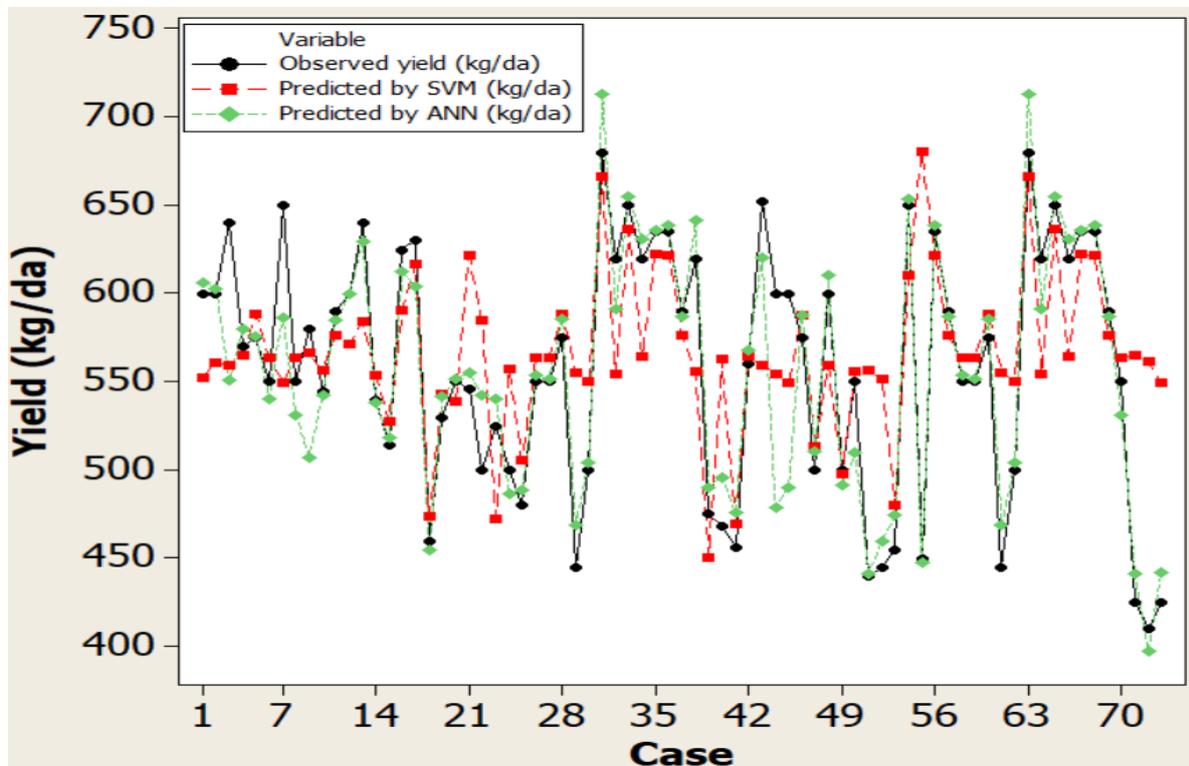

Figure 6. Observed yield values, yield values estimated with SVM and estimated yield values with ANN

In order to increase the accuracy of the model in the future, it is envisaged to investigate other factors that affect efficiency and to create new models. In addition to this, we believe that more accurate results can be obtained by conducting research not only among the cotton producers in Diyarbakır but also among all nearby producers.

Conclusion

In this study, it was tried to predict cotton production with factors such as Cotton area (da), Fertilizer (kg da-1), Tractor usage (h da-1) and Irrigation frequency using ANN and SVM model. The neural network model created with these factors can predict cotton production with an 84% accuracy.

The results of this study demonstrated the capacity of ANN models to better forecast cotton production applying different factors than a support vector machine model. In the future, it will be possible to develop high-accuracy ANN models with the addition of environmental and breeding factors affecting production and forbye, by using more business data.

The results of this study are the running start in developing suitable methods for estimating cotton production for the Diyarbakır region. Estimation models of different agricultural products can be created by using ANN method.

References

Anonym 2020a, Diyarbakır ve Şanlıurfa illerinde pamuk sektörü Envanteri, Access link: https://www.karacadag.gov.tr/Dokuman/Dosya/www.karacadag.org.tr_8_WH3D93PC_Diyarbakir_ve_Sanliurfa_illerinde_pamuk_sektoru_envanterinin_hazirlanmasi_projesi.pdf, Access date: 21.01.2020

Anonym, 2020b, Cotton production report, April 2019, Access Link: https://ticaret.gov.tr/data/5d41e59913b87639ac9e02e8/d0e2b9c79234684ad29baf256a0e7dce.pdf , Access date: 02.05.2020

Anonym 2020c, Cotton report, 2018 Access Link: http://www.zmo.org.tr/genel/bizden_detay.php?kod=30467&tipi=17&sube=0, Access date: 21.01.2020

Anonym 2020d, Cotton, Access link, https://arastirma.tarimorman.gov.tr/tepge/Belgeler/PDF%20Tar%C4%B1m%20%C3%9Cr%C3%BCnleri%20Piyasalar%C4%B1/2019Ocak%20Tar%C4%B1m%20%C3%9Cr%C3%BCnleri%20Raporu/2019-Ocak%20Pamuk.pdf, Access date: 21.01.2020

Anonym, 2020e. Diyarbakır ili coğrafyası. http://www.diyarbakirkulturturizm.gov.tr/TR-56881/cografya.html, Erişim tarihi: 01.05.2020 (In Turkish)

Bozkurt, Y., Aydogan, T., Tuzun, C.G., Mikail, N., Varban, S., Dogan, C., Tatlı, M. (2015). Some applications of artificial neural networks used beef cattle production, 4th International Congress New Perspectives and Challenges of Sustainable Livestock Production, 07-09 EKİM 2015, Belgrade, Sırbistan.

Çiçek, A., Erkan, O. (1996). Tarım Ekonomisinde Araştırma ve Örnekleme Yöntemleri. Gaziosmanpaşa Üniversitesi, Ziraat Fakültesi Yayınları No: 12, Ders Notları Serisi No: 6, Tokat.

Farooq, A., Sarwar, M. I., Ashraf, M. A., Iqbal, D., Hussain, A., & Malik, S. (2018). Predicting Cotton Fibre Maturity by Using Artificial Neural Network. *Autex Research Journal*. https://doi.org/10.1515/aut-2018-0024

Karademir, E , Karademir, C , Ekinci, R , Sevilmis, U. (2015). İleri Generasyondaki Pamuk (Gossypium hirsutum L.) Hatlarında Verim ve Lif Kalite Özelliklerinin Belirlenmesi. Türkiye Tarımsal Araştırmalar Dergisi, 2 (2) , 100-107 . DOI: 10.19159/tutad.60964 (In Turkish)

Khoshnevisan, B., Rafiee, S., Omid, M., Mousazadeh, H., & Sefeedpari, P. (2013a). Prognostication of environmental indices in potato production using artificial neural networks. *Journal of Cleaner Production*. https://doi.org/10.1016/j.jclepro.2013.03.028

Khoshnevisan, B., Rafiee, S., Omid, M., Yousefi, M., & Movahedi, M. (2013b). Modeling of energy consumption and GHG (greenhouse gas) emissions in wheat production in Esfahan province of Iran using artificial neural networks. *Energy*. https://doi.org/10.1016/j.energy.2013.01.028

Khoshroo, A., Emrouznejad, A., Ghaffarizadeh, A., Kasraei, M., & Omid, M. (2018). Sensitivity

analysis of energy inputs in crop production using artificial neural networks. *Journal of Cleaner Production*. https://doi.org/10.1016/j.jclepro.2018.05.249

Konishi, S. (2014). Introduction to multivariate analysis: Linear and nonlinear modeling. In *Introduction to Multivar. Analysis: Linear and Nonlinear Modeling*. https://doi.org/10.1201/b17077

Mammadova, N., & Keskin, I. (2013). Application of the support vector machine to predict subclinical mastitis in dairy cattle. *The Scientific World Journal*, *2013*, 603897. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3886278&tool=pmcentrez&rendertype=abstract

Mammadova, N. M., & Keskin, I. (2015). Application of neural network and adaptive neuro-fuzzy inference system to predict subclinical mastitis in dairy cattle. *Indian Journal of Animal Research*, *49*(5). https://doi.org/10.18805/ijar.5581

Mathworks (2009). MATLAB - Mathworks - MATLAB & Simulink. http://doi.org/2016-11-26

Mikail, N., Altay, Y., Keskin, İ. (2013). A Sample Model Prediction of 305-Day Milk Yield of Holstein Cows Using Artificial Neural Networks, VIth International Balkan Animal Conference, BALNIMALCON, Tekirdağ, Turkey, 03-05 EKİM, 2013.

Mikail, N., Keskin, İ., Altay, Y. (2014). Siyah Alaca ineklerin süt verimi tahmininde yapay sinir ağları ve destek vektör makineleri yöntemlerinin kullanımı, Uluslararası Mezopotamya Tarım Kongresi, 22-25 EYLÜL 2014, Diyarbakır, Türkiye.

Mikail, N., Keskin, İ. (2015). Application of neural network and adaptive neuro-fuzzy inference system to predict subclinical mastitis in dairy cattle. Indian J. Anim. Res., 49 (5) 2015 : 671-679.

Mikail, N., Keskin, İ., Altay, Y., Dağ, B. (2016). A sample model prediction of milk yield in Akkaraman ewes using artificial neural networks. International Human and Nature Sciences: Problems and Solution Seeking Congress, 07-09 EKİM 2016, Bosna Hersek.

Nabavi-Pelesaraei, A., Rafiee, S., Hosseinzadeh-Bandbafha, H., & Shamshirband, S. (2016). Modeling energy consumption and greenhouse gas emissions for kiwifruit production using artificial neural networks. *Journal of Cleaner Production*. https://doi.org/10.1016/j.jclepro.2016.05.188

Nguyen, H. (2019). Support vector regression approach with different kernel functions for predicting blast-induced ground vibration: a case study in an open-pit coal mine of Vietnam. *SN Applied Sciences*, *1*(4). https://doi.org/10.1007/s42452-019-0295-9

Rostami, S., Choobin, S., Samani, B. H., Esmaeili, Z., & Zareiforoush, H. (2017). *Analysis and Modeling of Yield, CO2 Emissions, and Energy for Basil Production in Iran using Artificial Neural Networks*.

Safa, M., & Samarasinghe, S. (2011). Determination and modelling of energy consumption in wheat production using neural networks: "A case study in Canterbury province, New Zealand." *Energy*. https://doi.org/10.1016/j.energy.2011.06.016

Safa, M., Samarasinghe, S., & Nejat, M. (2015). Prediction of wheat production using artificial neural networks and investigating indirect factors affecting it: Case study in canterbury province, New Zealand. *Journal of Agricultural Science and Technology*.

Saltuk, B., Mikail, N. (2019). Prediction of indoor temperature in a greenhouse: Siirt sample. Fresen. Environ. Bull. 28(4A), 3577-3585.

Samarasinghe, S. (2006). Neural Networks for Applied Sciences and Engineering. In *Neural Networks for Applied Sciences and Engineering*. https://doi.org/10.1201/9781420013061

Spiegel, M. R., Schiller, J. J., & Srinivasan, R. A. (2009). Probability and statistics. In *Schaum's outlines*.

Taki, M., Ajabshirchi, Y., & Mahmoudi, A. (2012). Prediction of output energy for wheat production using artificial neural networks in Esfahan province of Iran. *International Journal of Agricultural Technology, 8, 1229-1242.*

Taki, M., Rohani, A., Soheili-Fard, F., & Abdeshahi, A. (2016). Assessment of energy consumption and modeling of output energy for wheat production by neural network (MLP and RBF) and Gaussian process regression (GPR) models. *Journal of Cleaner Production*. https://doi.org/10.1016/j.jclepro.2017.11.107

Vapnik, V. (1996). The Nature of Statistical Learning Theory. Springer, New York.

Vapnik, V. (1998). Statistical Learning Theory. Wiley, New York.

Willmott, C.J. and Matsuura, K. (2005). Advantages of the Mean Absolute Error (MAE) over the Root Mean Square Error (RMSE) in Assessing

Average Model Performance. Climate Research, 30, 79-82. http://dx.doi.org/10.3354/cr030079

Yamane, T. (1967). Statistics: And Introductory Analysis, 2nd Ed., : In *Scottish Journal of Arts, Social Sciences and Scientific Studies*.

Yingli, L. V., Le, Q. T., Bui, H. B., Bui, X. N., Nguyen, H., Nguyen-Thoi, T., Dou, J., & Song, X. (2020). A comparative study of different machine learning algorithms in predicting the content of ilmenite in titanium placer. *Applied Sciences (Switzerland)*, *10*(2). https://doi.org/10.3390/app10020635