

Deep Neural Networks and an Application in Health Sciences

Derin Sinir Ağları ve Sağlık Bilimlerinde Bir Uygulama

Sadi Elasan¹

¹ Yüzüncü Yıl University, Faculty of Medicine, Department of Biostatistics, Van, Turkey

Abstract

Objective: Because there is more than one hidden layer between the input and output layers in the neural network algorithm, it is called "Deep Neural Networks". In the study, the Deep Neural Networks algorithm; different input (number of layers, epoch, error rate) and evaluation of the performance of the model being practiced an application is intended.

Materials and Methods: The most important feature that distinguishes the deep neural network method from the classical neural networks is the number of layers that provide good results in complex problems. Wart treatment results of patients who used immunotherapy were used as data set in the study.

Results: According to this; the simple-layer (1 hidden layer) artificial neural network model was classified with 87.5% overall accuracy and 29.74% MAPE, whereas the deep neural network model was classified with 99.8% overall accuracy and 25.19% MAPE ratio. The study showed that the model of deep neural networks had a higher accuracy rate.

Conclusion: As a result of the application performed in this study; it is seen that the multi-layered (deep) neural network model provides classification with higher Accuracy and lower error rate than the single layer (classical) neural network model. In other words, according to the results of this study; it has been found that the deep neural network model has a better (optimum) classification rate than the classical neural network model.

Key Words: Deep neural networks; immunotherapy; multilayer network

Özet

Amaç: Yapay sinir ağları algoritmasında giriş ve çıkış katmanları arasında birden fazla gizli katman bulunduğu için, "Derin Sinir Ağları" (Deep Neural Network) olarak adlandırılır. Klasik bir yapay sinir ağına bulunan nöronların birbirleriyle ilişkileri yoktur ve bilgi ancak giriş katmanından çıkış katmanına doğru aktarılır. Derin sinir ağlarında ise artarda gelen iki katmanda nöronlar birbirlerini çeşitli aktivasyon değerleriyle etkilemektedir. Her katmanın modele etkisi ve dolayısıyla her katmandaki nöronun modele etkisi bulunmaktadır. Çalışmada, Derin Sinir Ağları algoritması; farklı girdi (katman sayısı, döngü, hata oranı) ve uygulamalı modelin performansının değerlendirilmesi amaçlanmaktadır.

Gereç ve Yöntem: Derin sinir ağı yöntemini klasik sinir ağlarından ayıran en önemli özellik, karmaşık problemlerde iyi sonuçlar veren katman sayısıdır. Çalışmada veri seti olarak immünoterapi kullanan hastaların siğil tedavisi sonuçları verisi kullanıldı.

Bulgular ve Sonuç: Basit katmanlı (bir gizli katman) yapay sinir ağı modelinde %87.5 genel doğruluk ve %29.74 MAPE ile sınıflandırılırken, derin sinir ağı modeli %99.8 genel doğruluk ve %25.19 MAPE oranı ile sınıflandırılmıştır. Çalışma, derin sinir ağları modelinin daha yüksek doğruluk oranına sahip olduğunu göstermiştir.

Anahtar Kelimeler: Derin sinir ağları; İmmünoterapi; Çok katmanlı ağı

Introduction

Because there is more than one hidden layer between the input and output layers in the neural network algorithm, it is called "Deep Neural Networks". In a classical neural network, neurons in a layer do not have interrelationships and information is only transferred from the previous layer to the next layer or exit. In the deep neural networks, the neurons affect each other with various activation values from two successive layers. The effect of each layer on the model and thus the neuron in each layer affects the model. In the study, the Deep Neural Networks algorithm; different input (number of layers, epoch, error rate) and evaluation of the performance of

the model being practiced an application is intended. The most important feature that distinguishes the deep neural network method from the classical neural networks is the number of layers that provide good results in complex problems. Warts are benign skin formations originating from a particular group of viruses known as human papillomaviruses. Although there is no definite answer to the question of why a wart occurs, it can be said that it occurs more frequently in people with weak immune systems. There are different methods and solutions used in the treatment of warts. After treatment, the virus becomes passive in some people and the wart does not recur. In some individuals, there is the

possibility of recurrence due to the weakness of the immune system. For this reason, immune-enhancing measures are frequently used after treatment. According to the literature review, it is stated that immunotherapy is effective in the treatment of warts. Immunotherapy is an effective treatment for resistant warts and prevents potential problems. Therefore, there is a need for safe, inexpensive, effective and simple immunotherapeutic agents for the treatment of warts (1-2-3-4). Wart treatment results of patients who used immunotherapy were used as data set in this study. This data was modelled using the “Deep Neural Networks” method. Thus, the main variables that affect the success of immunotherapy have been tried to be determined.

Materials and Methods

In this study, the data set given descriptive statistics in Table 1 was used to make an application of deep neural networks. The application material of this study was obtained from UCI (5) from the free-access data site (6). Wart treatment results of the patients using immunotherapy were used as the dependent variable. For data analysis, “Weka (ver: 3.9) and SPSS (ver: 23)” statistical package programs were used. In this statistical program, from the “multi-layer sensor” models; by using activation functions "Hyperbolic Tangent, Sigmoid and Softmax" the application was performed.

Table 1: Descriptive statistics

		Mean	SD
Age (Year)		31.044	12.24
Time Elapsed Before Treatment (M)		7.231	3.10
Num of Warts		6.144	4.21
Surface Area* (mm ²)		95.7001	136.61
Diameter of Initial Test (mm)		14.333	17.22
		n	%
Sex	Man	41	45.6
	Female	49	54.4
Type of Warts	1 Common	47	52.2
	2 Plantar	22	24.4
	3 Both	21	23.3
Immunotherapy	Successful	71	78.9
	Unsuccessful	19	21.1

SD: Standard Deviation

In the hidden layers in the model, forward calculations and backward error propagation were performed. This function is quite useful particularly for the activation of nerve cells in an artificial neural network. In this model, the neural

network can be installed by the user or automatically by the program. Network model parameters can be monitored and changed during the training period. In this study, the performance of the model was evaluated by making an application of a Deep Neural Networks algorithm with different input (Epoch: 500, Learning Rate: 0.3, Hidden layer number: 0-6) alternatives.

Deep (multilayer) neural networks: Artificial neural networks are formed by a mathematical model inspired by the structure and function of biological neural networks. Artificial neural networks are often used to model the relationship between input and output. Deep neural networks algorithm can be considered as a structurally more complex form of artificial neural networks. Because there is more than one hidden layer between the input and output layers the artificial neural network algorithm is called “Deep Neural Networks”. Neurons in a conventional artificial neural network do not have a relationship with each other and information is only transferred from the input layer to the output layer. In deep neural networks, neurons affect each other with various activation values in two consecutive layers. Each layer affects the model and therefore the neuron in each layer affects the model. Deep learning consists of a multi-layered structure that is assembled to learn the model. The most important feature that distinguishes the method of deep neural networks from classical artificial neural networks is the number of layers (depth) that provides good results in complex problems. This study, it is aimed to evaluate the performance of the model by making an application of the Deep Neural Networks algorithm with different input (hidden layer and neuron numbers) alternatives. The deep learning algorithm is slightly different from machine learning algorithms. The most important of these differences is the use of large amounts of data in the deep learning algorithms and at the same time, they need higher hardware. (7-8). Deep neural network architecture consists of an input layer, multiple hidden layers and output layers. Input Layer: It is the layer in which the attributes of a sample that comes as input and requested to be learned are given as entry. In the input layer, there must be an input neuron equal to the number of attributes of the samples to be learned. Output Layer: It is the element where the class information or label value of the samples to be learned in the network is calculated as output.

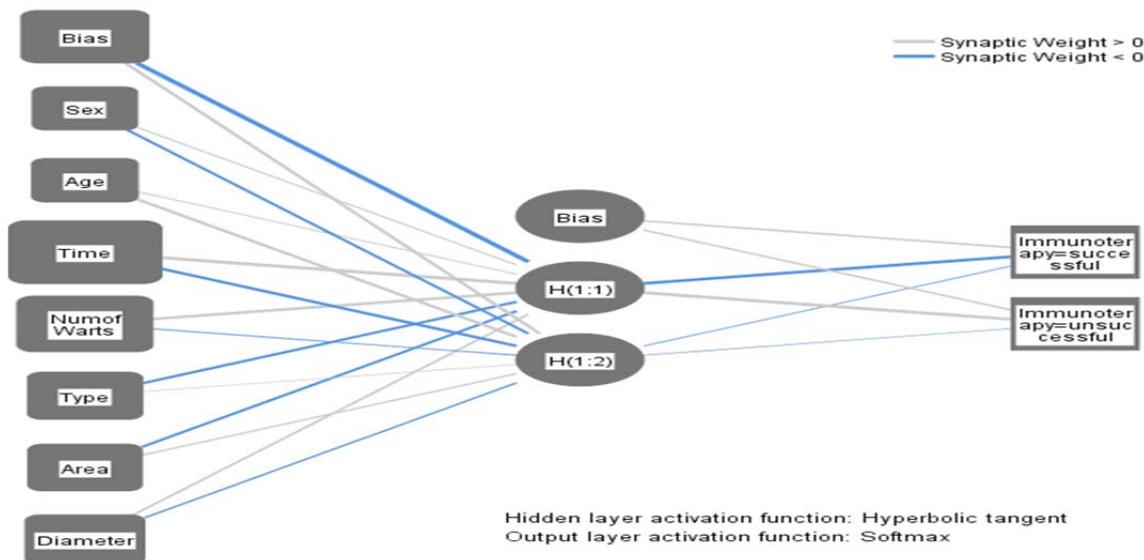


Figure 1. A deep neural network model for two hidden layers

Table 2: Classification performances of Artificial Neural Networks according to Layers

Performance Criteria*	Hidden Layers					
	1	2	3	4	5	6
Correctly Classified (%)	74.19	80.65	83.87	83.87	83.87	83.87
Kappa statistic %	27.91	38.41	00.00	00.00	00.00	00.00
Mean absolute error % (MAPE)	29.74	25.19	31.67	31.67	31.67	31.67
Relative absolute error (%)	90.70	76.83	96.60	96.60	96.60	96.60
Overall Accuracy (%)**	87.50	99.8	84.40	85.70	88.00	80.00
Num of Neurons: 5 Num of Epoch: 500	Learning Rate: 0.3		Momentum: 0.2			

Table 3: Independent variable importance for hidden layers

	Hidden Layers						
	1	2	3	4	5	6	7
Sex	0.8%	5.9%	26.3%	7.9%	12.4%	24.1%	15.9%
Age (Year)	16.4%	7.7%	69.6%	16.9%	10.6%	43.0%	64.8%
Time Elapsed Before Treatment (M)	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Num of Warts	22.0%	66.0%	18.1%	22.0%	15.3%	33.6%	62.7%
Type of Warts	8.1%	25.1%	8.7%	9.0%	10.9%	6.2%	45.6%
Surface Area (mm ²)	14.5%	23.0%	26.7%	29.9%	48.5%	50.1%	65.0%
Diameter of Initial Test (mm)	17.7%	33.6%	33.8%	30.3%	20.4%	68.0%	53.8%

Hidden Layers: These are the layers between the input layer and the output layers. The number of layers and the number of neurons on them can vary from problem to problem. In these layers, forward calculations and reverse error propagation are performed. In the solution of complex problems, the number of layers and the number of neurons in the layers are generally higher. Each layer in the network affects the model, and the neuron in each layer affects the layer, hence on

the model. Neurons in two consecutive layers affect each other with various activation values. The number of neurons in the layers indirectly affects the performance of the system (9-10).The number of Layers and Number of Neurons in Hidden Layers: It is observed that the model learns better as the number of layers increases in complex problems. In the multilayer artificial neural networks, it was found that firstly the general lines of the data are learned, and in later

layers, the other features of the data are gradually being learned. The number of neurons in the layers indicates the number of information stored in memory (11).

Learning rule of multilayer network: “Learning Rule of a Multilayer Network” is a generalized

version of “Delta Learning Rule” based on the least-squares method. It is therefore also known as the "Generalized Delta Rule". “Generalized Delta Rule” consists of two stages (12).

Table 4: The effect of variables on the success of immunotherapy

		Immunotherapy	
		Successful	Unsuccessful
		Mean	Mean
Age (Year)		29.86	35.47
Time Elapsed Before Treatment (M)		6.65	9.38
Num of Warts		6.04	6.53
Surface Area* (mm ²)		98.75	84.32
Diameter of Initial Test (mm)		14.06	15.37
		%	%
Sex	Male	45.1	47.4
	Female	54.9	52.6
Type of Warts	1 Common	50.7	57.9
	2 Plantar	23.9	26.3
	3 Both	25.4	15.8

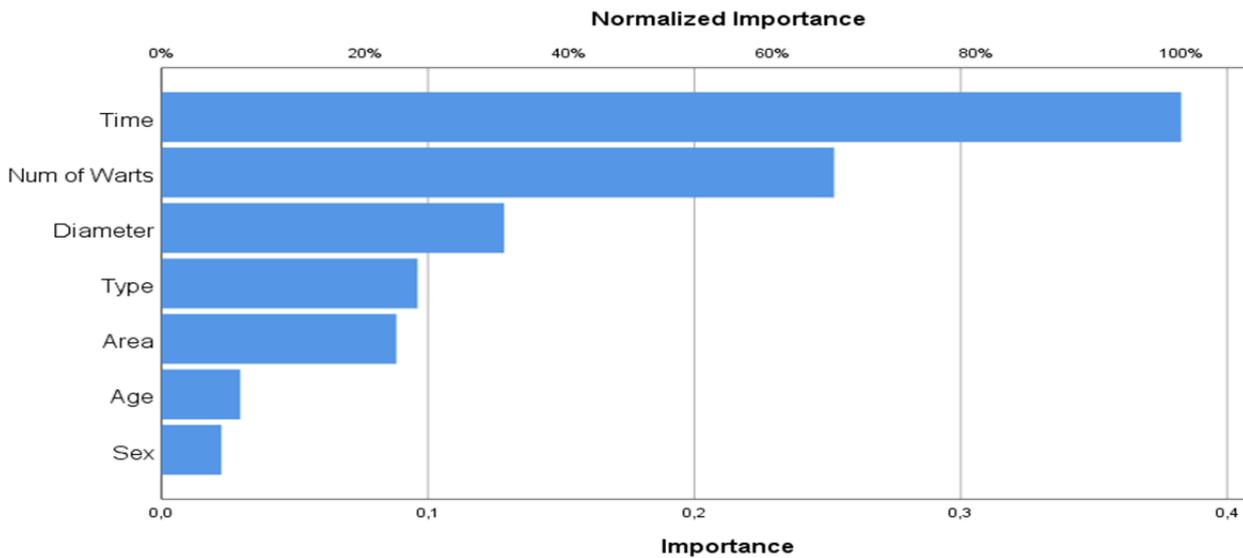


Figure 2.Independent variables importance for two hidden layers

Feed forward calculation:At this stage, information processing begins by showing a sample in the training set to the network in the input layer. Incoming inputs are sent to the intermediate layer without any changes (12-13).The output of the k. element in the input layer: $C_k^i = G_k$

Each processing element in the intermediate layer receives information from all process elements in

the input layer under the influence of the link weights (A1, A2,...) The net (N) input to the process elements in the intermediate layer is calculated using the following formula.

$$N_j^a = \sum_{k=1}^n A_{kj} C_k^i$$

N: each neuron; n: the number of connections of a neuron; $A_{kj}G_k^i$: connection of each input to the related neuron; A_{kj} k. input element j. the weight value of the connection connecting to the intermediate layer element; j. The output of the intermediate layer element. In the widely used multi-layer sensor model, the ‘‘Sigmoid function’’ is commonly used. The output if the sigmoid function is used:

$$C_j^a = \frac{1}{1 + e^{-(N_j^a + \beta_j^a)}}$$

Here β_j , represents the weight of the threshold element connected to the j. element in the intermediate layer. The output of this threshold unit is constant and is equal to 1. During training, the network determines this value (12-13).

Back propagation calculation: The output produced by the network with the help of input presented to the network is compared with the expected outputs of the network. The difference between them is considered an error. The aim is to reduce this error. Distributing this error to the weight values of the web provides reducing an error in the next iteration (12-14). Error for processing element m. in output layer: $E_m = B_m + C_m$ Sum of errors for error (TH) that occurs for the output layer: $TH = 1/2(\Sigma_m^2)$

To minimize the total error, it must be distributed to the processing elements that cause this error.

Changing weights between the intermediate layer and output layer: If the amount of change in the weight of the connection connecting the process element j. to the process element m. in the output layer is called ΔA^a ; The amount of change of weight at any t-time is calculated as follows (12):

$$\Delta A_{jm}^a(t) = \lambda \delta_m C_j^a + \alpha \Delta A_{jm}^a(t-1)$$

Here λ represents the learning coefficient, α represents the momentum coefficient, and δ_m represents the error of the output unit m. (12-14).

Results

Neurons in a conventional artificial neural network have no relationship with each other and information is only transferred from the input layer to the output layer. In deep neural networks, neurons affect each other with various activation values in two consecutive layers. Each layer affects the model and hence the neuron in each layer affects the model (Figure 1). The ‘‘two-layer neural network model’’ which gives the highest overall accuracy as a result of the application is given

above (Figure 1). Table 2. Classification performances of Artificial Neural Networks according to Layers As a result of the application, while the number of layers that gives the highest accuracy rate in the deep neural network model was determined as two (2). The classical (single-layer) artificial neural network model was classified with 87.5% overall accuracy and 29.74% MAPE. Thus, in the case of ‘‘2 layers’’ in the network, the best (optimum) classification was achieved with 99.8% General Accuracy and 25.19% MAPE. In case the of ‘‘number of layers 3’’ was classified with %84,4 Accuracy and MAPE %31,67. Similarly, if the number of layers is entered as 4 or more, the classification results in the table are obtained. Table 3. Independent variable importance for hidden layers at the end of the study, the significance of the independent variables according to the number of layers and their effects on the success of immunotherapy were given. According to this; in the case where the neural network is bi-layer, the most effective variables on immunotherapy are respectively determined as; ‘‘Time Elapsed Before Treatment, Number of Warts, Diameter of Initial Test, Type of Warts, Surface Area, Age and Sex’’ (Table 3, Table 4 and Figure 2). When the effect of the variables on the success of immunotherapy is examined; Age, Time Elapsed Before Treatment, Numerous Warts and Diameter of Initial Test increases the success of treatment decreases. However, success increases as the ‘‘surface area of the biggest wart’’ increases. Also, treatment is more successful in women than in men (Table 4).

Discussion

The most important feature that distinguishes the method of deep neural networks from classical artificial neural networks is the number of layers (depth) that provides good results in complex problems. It has been observed that the model learns better as the number of layers that allows it to give good results in complex problems increases. In the multilayer artificial neural networks, it was found that firstly the general lines of the data are learned, and in later layers, the other features of the data are gradually being learned. Although there is no general method for the number of layers, some principles have been proposed. For example, firstly one may start with a single-layer or two-layer network and focus on finding suitable parameters. Then, the number of layers can be increased gradually. As the number of layers’ increases, the back propagation effect can reach the first layers less. The number of

neuron in the layers indicates the number of information stored in memory. The higher the number of neurons, the higher the memory requirement and computational time. However, a small number of neurons causes under fitting. Another feature is the different distribution of the number of neurons between the layers; that is to say, while more neurons are used in the first layers, continuation by reducing it in the following layers creates a regularization effect (Brownlee, 2016; Çarkacı, 2018). As a result of the application performed in this study; it is seen that the multi-layered (deep) neural network model provides classification with higher Accuracy and lower error rate than the single layer (classical) neural network model. In other words, according to the results of this study; it has been found that the deep neural network model has a better (optimum) classification rate than the classical neural network model.

Ethical Approval: This study data consists of data obtained from the UCI University (Irvine) open-source data and previously approved by the ethics committee (5).

Conflict of Interest: There is no conflict of interest in this study

Financial Support: No financial support was used in this study.

Author Contributions: The design of this study, source review, statistical analysis and modeling were done by Sadi Elasan.

References

1. Naylor MF, Neldner KH, Yarbrough GK, Rosio TJ, Iriando M, Yeary J. Contact immunotherapy of resistant warts. *J Am Acad Dermatol* 1988; 19(4):679-683.
2. Zur Hausen H, de Villiers EM. Human papillomaviruses. *Annu Rev Microbiol* 1994; 48(1):427-447.
3. Chandrashekar L. Intralesional immunotherapy for the management of warts. *Indian J Dermatol Venereol Leprol* 2011; 77(3):261-262.
4. Khozeimeh F, Alizadehsani R, Roshanzamir M, Khosravi A, Layegh P, Nahavandi S. An expert system for selecting wart treatment method. *Comput Biol Med* 2017; 81;2(1):167-175.
5. UCI. 2019. University of California, Irvine. <https://archive.ics.uci.edu/ml/datasets/Immunotherapy+Dataset> Accessed: 15.06.2019.
6. Khozeimeh F, Jabbari Azad F, Mahboubi Oskouei Y, Jafari M, Tehranian S, Alizadehsani R, et al. Intralesional immunotherapy compared to cryotherapy in the treatment of warts. *Int J Dermatol* 2017; 56(4):474-478.
7. Rende FŞ, Bütün G, Karahan Ş. Derin Öğrenme Algoritmalarında Model Testleri: Derin Testler. 10. Ulusal Yazıl. Müh. Semp 24-26 Ekim 2016; 54-59, Çanakkale.
8. Wang Y, Mao H, Yi Z. Protein secondary structure prediction by using deep learning method. *Knowledge-Based Systems* 2017; 118:115-123.
9. LeCun Y, Bengio Y, Hinton G. Deep learning, *Nature* 2015; 521:436-444.
10. Brownlee J. 2016. How to Grid Search Hyperparameters for Deep Learning Models in Python with Keras, <https://machinelearningmastery.com/grid-search-hyperparameters-deep-learning-models-python-keras/>, Accessed: 21.04.2019.
11. Çarkacı N. 2018. Derin Öğrenme Uygulamalarında En Sık kullanılan Hiper parametreler. <https://medium.com/deep-learning-turkiye/derin-ogrenme-uygulamalarinda-en-sik-kullanilan-hiper-parametreler-ece8e9125c4> Accessed: 19.04.2019.
12. Schmidhuber J. Deep learning in neural networks: An overview. *Neur Netw* 2015; 61: 85-117.
13. Güzel Y. 2018. Yapay Zeka Ders Notları 05. Çok Katmanlı Yapay Sinir Ağı. <https://medium.com/@yasinguzel/yapay-zeka-ders-notlar%C4%B1-5-%C3%A7ok-katmanl%C4%B1-yapay-sinir-a%C4%9F%C4%B1-4c6af075e1fe> Accessed: 02.05.2019.
14. Güzel K. 2018. Geri Yayılımlı Çok Katmanlı Yapay Sinir Ağları-1. <https://kadirguzel.medium.com/geri-yay%C4%B1ml%C4%B1ml%C4%B1-%C3%A7ok-katmanl%C4%B1-yapay-sinir-a%C4%9Flar%C4%B1-1-47daa3856247> Accessed: 15.04.2019.