



## PERFORMANCE COMPARISON OF NEURAL NETWORK-BASED MODELS IN THE CLASSIFICATION OF POLYCYSTIC OVARY SYNDROME DISEASE

Zeynep KÜÇÜKAKÇALI<sup>1</sup>, Fatma Hilal YAĞIN<sup>1</sup>, İpek BALIKÇI ÇİÇEK<sup>1\*</sup>

<sup>1</sup>Inonu University, Faculty of Medicine, Department of Biostatistics and Medical Informatics, 44000, Malatya, Türkiye

**Abstract:** In this study, it was aimed to compare the performances of the above mentioned ANN, MLP and deep learning methods to determine polycystic ovary syndrome (PCOS) risk factors and predict PCOS diagnosis. In this study, the data set "Polycystic ovary syndrome" was used to determine PCOS risk factors and to compare the performances of ANN, MLP and deep learning methods for PCOS diagnosis prediction. The performance of the models was evaluated with accuracy, sensitivity, specificity, positive/negative predictive values. Factors associated with PCOS were estimated from the deep learning model that has the best performance. Accuracy, sensitivity, specificity, positive predictive value and negative predictive value obtained from the MLP method were 87.25%, 79.66%, 90.93%, 81.03%, and 90.19%. Accuracy, sensitivity, specificity, positive predictive value and negative predictive value obtained from the Neural Network method were 87.80%, 79.10%, 92.03%, 82.84%, and 90.05%. Accuracy, sensitivity, specificity, positive predictive value and negative predictive value obtained from the Deep Learning method were 89.09%, 81.92%, 92.58%, 84.30%, and 91.33%. According to the findings obtained from this study, the best classification result according to the performance metrics obtained from the artificial neural networks, MLP and deep learning methods used for the PCOS data set used in the study belongs to the deep learning method. As a result, PCOS was successfully classified in the light of the findings obtained from the study, and clinical findings were tried to be revealed by giving the risk factors associated with PCOS.

**Keywords:** PCOS, MLP, ANN, Deep learning, Classification

\*Corresponding author: Inonu University, Faculty of Medicine, Department of Biostatistics and Medical Informatics, 44000, Malatya, Türkiye

E-mail: ipek.balikci@inonu.edu.tr (İ. BALIKÇI ÇİÇEK)

Zeynep KÜÇÜKAKÇALI <https://orcid.org/0000-0001-7956-9272>

Fatma Hilal YAĞIN <https://orcid.org/0000-0002-9848-7958>

İpek BALIKÇI ÇİÇEK <https://orcid.org/0000-0002-3805-9214>

Received: July 15, 2022

Accepted: September 25, 2022

Published: January 01, 2023

**Cite as:** Küçükakçalı Z, Yağın FH, Balıkçı Çiçek İ. 2023. Performance comparison of neural network-based models in the classification of polycystic ovary syndrome disease. BSJ Health Sci, 6(1): 20-25.

### 1. Introduction

Polycystic ovary syndrome (PCOS) is considered to be one of the most common endocrine and metabolic disorders affecting 9-18% of women of reproductive age, depending on the diagnostic criteria and the population studied. Cardiometabolic diseases such as insulin resistance, hyperinsulinemia and obesity, infertility and many psychological diseases are closely related with PCOS (Azziz, 2016; Steegers-Theunissen et al., 2020). PCOS can be detected by biochemical, clinical, and ultrasound approaches. Early detection and therapy have been shown to minimize the risk of PCOS. A woman diagnosed with PCOS poses a significant burden on the healthcare system due to the need for constant medical follow-up. Accurate diagnosis is essential to overcome this problem. However, the etiopathogenesis of PCOS, which has a chronic course, cannot be fully explained. This situation causes diagnostic complexity (Yüce et al., 2020).

Nowadays, machine learning (ML) methods are used by researchers as a non-invasive method for predicting diseases. ML methods are frequently used in PCOS diagnosis prediction in the literature. PCOS datasets such

as clinical, biochemical, medical history, patients' symptoms, and ultrasound images are used to create predictive models (Satish et al., 2020).

Artificial neural networks (ANN), one of the ML methods, are computational structures that imitate the cognitive learning process of the human brain. ANN consists of process elements, each of which has its own memory, connected by means of heavy links. Unlike classical calculation methods, ANN is quite successful in its ability to solve nonlinear problems such as estimation, optimization, and recognition and control (El-Mahelawi et al., 2020).

Multilayer perceptions (MLP), a type of artificial neural networks, are frequently used in disease prediction. In MLP, neurons are organized and arranged in layers and there are three layers. From these layers, the input layer contains information about the problem to be solved. The other layer is the output layer from which the processed information in the classification network is output. The layer between the input and output layers is called the hidden layer (Karabulut and Sarac, 2020).

Deep learning can be considered as the structurally more developed version of artificial neural networks. Deep



learning is frequently used in many areas such as image analysis, voice analysis, robotics, gene and protein analysis, cancer diagnosis and virtual reality. The most important reason why it is used in such a widespread area is the high accuracy it obtains in solving problems. In this method, there are more than one hidden layer between the input and output layers. The method uses many layers of nonlinear processing units for feature extraction and conversion. Each successive layer takes the output from the previous layer as input (Wang et al., 2020).

In this study, it was aimed to compare the performances of the above mentioned ANN, MLP and deep learning methods to determine PCOS risk factors and predict PCOS diagnosis.

## 2. Materials and Methods

### 2.1. Dataset

In this study, the data set named "Polycystic ovary syndrome" was obtained from <https://www.kaggle.com/prasoonkottarathil/polycystic-ovary-syndrome-pcos> address to determine PCOS risk factors and to compare the performances of ANN, MLP and deep learning methods for PCOS diagnosis prediction. In the PCOS data set, there are a total of 541 patients, 364 (67.3%) no and 177 (32.7%) yes.

### 2.2. Artificial Neural Networks (ANN)

Artificial Neural Networks (ANN) is a model inspired by living organisms with its biological neural structure. ANN is a technique used to imitation a goal-oriented brain structure by training the network between input and output based on solving a problem, predicting, and determining levels of significance. The artificial neural network's technical task is to convert the varied data acquired by the input layer from the outside world into information that can be used as output by the network's processing and learning. Most artificial neural networks are completely connected from one layer to the next, and these connections have weight values. As the weight values increase, the effect of one unit on another increases similar to the human brain. As the data passes through these units, the artificial neural network learns more about the data. On the other side of the network are output units, the output units are where the network responds to the data given to it (Öztemel, 2003). ANN can provide nonlinear modeling without any prior knowledge between input and output variables, without any assumptions (Aggarwal, 2018). Artificial neural networks are the preferred method for many predictive data mining applications due to their efficiency, versatility, and ease of use. Predictive neural networks are particularly useful in applications that have complicated mechanisms. In recent years, there has been a surge of interest in using neural networks to solve problems that cannot be solved using traditional methods, and ANNs have been successfully used in a variety of medical applications (Orhan et al., 2010).

### 2.3. Multilayer Perceptions (MLP)

In the case of non-linear mapping, the MLP is arguably the most popular ANN. The input nodes receive the data values and pass them on to the first hidden layer nodes. Each one receives the input from all input nodes, adds a bias to the total, and sends the results through a non-linear transformation such the sigmoid transfer function after multiplying each input value by weight. This serves as the input for either the second hidden layer or the output layer, which works in the same way as the hidden layer. The network output is the result of each output node's transformed output. A training algorithm such as backpropagation, cascade correlation, or conjugate gradient must be used to train the network. The primary goal of training patterns is to reduce global error. By changing the weights and biases, every training algorithm aims to reduce this global error (Kashaninejad et al., 2009; Ayşe and Berberler, 2017).

### 2.4. Deep Learning

Deep learning is a field of research that deals with computer systems analyzing a problem based on existing data and generating outputs to solve the problem. It was inspired by the way the human brain operates. With the advancement of technology and breakthroughs in the field of computing, its popularity has risen dramatically in recent years (Esteva et al., 2019). Deep learning is a machine learning method that is used to solve problems and perform behavior such as analysis, inference, observation, and learning using large quantities of data, according to the descriptions in the literature. They can be in various hierarchical systems, unlike conventional machine learning algorithms. Deep learning is a sub-field of artificial neural networks (ANN). It's a type of ANN that uses nonlinear transformations to get a specific output value from raw data. Using the backpropagation algorithm, deep learning investigates the complex structure of multidimensional data sets. It accomplishes this by comparing the values of the parameters measured in each layer to the values obtained in the previous layer and determining the appropriate adjustment. Multi-layered neural networks are the foundation of deep learning. Multi-layer neural networks have an input layer that represents the inputs, hidden layers that transform the information from the input layer into an output, and an output layer that converts the results from the last hidden layer into output values (Kayaalp and Süzen, 2018). The success rate in the fields of natural language processing, image processing, visual object identification, and drug discovery has increased significantly thanks to deep learning methods (LeCun et al., 2015).

### 2.5. Statistical Analysis

For the outcome variable of patients with PCOS or absence to be examined using the G\*Power 3.1 program within the scope of the current study, Type I error amount (alpha) 0.05, test power (1-beta) 0.8, effect size 0.25, distribution ratio to groups 0.5 and alternative hypothesis (H1) was two-sided, with the help of two independent samples t-test, it was calculated by

theoretical power analysis that 175 and 351 patients should be included in the groups. The study included patients that 177 with PCOS and 364 without PCOS. Quantitative data are expressed as median (minimum-maximum), and qualitative data as number (percentage). Conformity to normal distribution was evaluated by the Kolmogorov-Smirnov test. In terms of independent variables, whether there is a statistically significant difference between the "no" and "yes" groups, which are the categories of the dependent / target variable (Pcos(Y/N)), and whether there is a relationship, Mann-Whitney U test, Pearson chi-square test. It was examined using the chi-square test values of  $p < 0.05$  were considered statistically significant. IBM SPSS Statistics 26.0 package program was used for all analyzes. In the modeling phase of the study, 10-fold cross-validation was used. Accuracy, sensitivity, specificity, positive predictive value, and negative predictive value

were used as performance evaluation criteria. In addition, variable importance were calculated, which gives information about how much the input variables explain to the output variable.

### **3. Results**

Table 1 shows descriptive statistics for numerical independent variables in this study, while Table 2 shows descriptive statistics for qualitative independent variables. According to the results in Table 1; there is a statistically significant difference between the dependent target variable (PCOS (Y/N)) groups in terms of Age (yrs), Weight (kg), Pulse rate (bpm), Hb (g/dL), Cycle (R/I), Cycle length (days), Marraige Status (Yrs), FSH(mIU/mL), Hip (inch), Waist (inch), AMH (ng/mL), Follicle No. (L), Follicle No. (R), Avg. F size (L) (mm), Avg. F size (R) (mm), Endometrium (mm), BMI and Fsh/Lh variables ( $p < 0.005$ ).

**Table 1.** Descriptive statistics for quantitative independent variables

Variables	PCOS (Y/N)		p-value*
	No Median (min-maks)	Yes Median (min-maks)	
Age (yrs)	32 (20-48)	29 (21-47)	<0.001
Weight (kg)	58 (32-108)	62 (31-104)	<0.001
Height (cm)	156 (137-173.7)	158 (143-180)	0.160
Pulse rate (bpm)	72 (13-82)	72 (70-82)	0.004
RR (breaths/min)	18 (16-28)	20 (16-24)	0.281
Hb (g/dL)	11 (8.5-14.8)	11 (9.4-14)	0.025
Cycle (R/I)	2 (2-4)	4 (2-5)	<0.001
CL (days)	5 (0-12)	5 (2-12)	<0.001
MS (Yrs)	7 (0-25)	6 (1-30)	0.003
No of abortions	0 (0-5)	0 (0-3)	0.424
I beta-HCG(mIU/mL)	13.735 (1.3-32460.97)	70.53(1.92-30007)	0.068
II beta-HCG(mIU/mL)	1.99 (0.99-21084.21)	1.99 (1.65-25000)	0.774
FSH (mIU/mL)	5.01 (0.21-5052)	4.48 (1-65.4)	0.007
LH (mIU/mL)	2.305 (0.02-14.69)	2.22 (0.032-2018)	0.353
Hip (inch)	38 (26-48)	39 (26-48)	<0.001
Waist (inch)	34 (24-46)	35 (24-47)	<0.001
TSH (mIU/L)	2.165 (0.04-65)	2.31 (0.05-22.59)	0.715
AMH (ng/mL)	3.2 (0.16-26.8)	5.9 (0.1-66)	<0.001
PRL (ng/mL)	21.17 (0.4-128.24)	22.9 (3.64-111.74)	0.592
Vit D3 (ng/mL)	26.3 (9.01-90)	25.45 (0-6014.66)	0.230
PRG(ng/mL)	0.31 (0.11-85)	0.32 (0.047-1.1)	0.385
RBS(mg/dl)	96 (60-2259)	100 (70-350)	0.345
BP _Diastolic (mmHg)	80 (8-100)	80 (70-80)	0.470
Follicle No. (L)	4 (0-15)	10 (1-22)	<0.001
Follicle No. (R)	4 (0-16)	11 (1-20)	<0.001
BP _Systolic (mmHg)	110 (12-140)	110 (100-130)	0.948
Avg. F size (L) (mm)	15 (0-22)	16 (5-24)	0.009
Avg. F size (R) (mm)	15 (0-24)	16 (0.17-23)	0.026
Endometrium (mm)	8.3 (0-18)	8.9 (4.5-15)	0.005
BMI	23.62 (13.99-38.27)	25.15 (12.42-38.90)	<0.001
Fsh/Lh	2.36 (0.23-1372.83)	2.04 (0.00-327.00)	0.006
Waist/Hip	0.89 (0.78-0.98)	0.90 (0.76-0.98)	0.652

\*Mann Whitney U test, CL= cycle length, MS= marraige status, F= follicle

According to the results in Table 2; there is a statistically significant relationship between the dependent / target variable (PCOS (Y/N)) groups in terms of weight gain (Y/N), hair growth (Y/N), skin darkening (Y/N), hair loss (Y/N), pimples (Y/N) and fast food (Y/N) variables ( $p<0.05$ ).

The metrics for the classification performance of MLP, Neural Network and Deep Learning methods in the test phase are given in Table 3. Accuracy, sensitivity, specificity, positive predictive value and negative predictive value obtained from the MLP method were 87.25%, 79.66%, 90.93%, 81.03%, and 90.19%.

Accuracy, sensitivity, specificity, positive predictive value and negative predictive value obtained from the Neural Network method were 87.80%, 79.10%, 92.03%, 82.84%, and 90.05%. Accuracy, sensitivity, specificity, positive predictive value and negative predictive value obtained from the Deep Learning method were 89.09%, 81.92%, 92.58%, 84.30%, and 91.33%.

According to the Deep Learning method, the five most important risk factors that may be associated with Pcos were obtained as Follicle No. (L), Avg. F size (L) (mm), Cycle (R/I), Vit D3 (ng/mL), and Hip (inch) variables.

**Table 2.** Descriptive statistics for qualitative independent variables

Variables	PCOS (Y/N)		p-value*
	No	Yes	
Blood Group	A+	74 (% 20.3)	0.957
	A-	9 (% 2.5)	
	B+	93 (% 25.5)	
	B-	10 (% 2.7)	
	O+	140 (% 38.5)	
	O-	11 (% 3.0)	
	AB+	26 (% 7.1)	
	AB-	1 (% 0.3)	
Pregnant(Y/N)	0	222 (% 61.0)	0.521
	1	142 (% 39.0)	
Weight gain (Y/N)	0	281 (% 77.2)	<0.001
	1	83 (% 22.8)	
Hair growth (Y/N)	0	317 (% 87.1)	<0.001
	1	47 (% 12.9)	
Skin darkening (Y/N)	0	308 (% 84.6)	<0.001
	1	56 (% 15.4)	
Hair loss (Y/N)	0	221 (% 60.7)	<0.001
	1	143 (% 39.3)	
Pimples (Y/N)	0	222 (% 61.0)	<0.001
	1	142 (% 39.0)	
Fast food (Y/N)	0	224 (% 61.7)	<0.001
	1	139 (% 38.3)	
Reg.Exercise(Y/N)	0	281 (% 77.2)	0.129
	1	83 (% 22.8)	

**Table 3.** Values for the metrics of the classification performance of MLP, Neural Network and Deep Learning methods

Methods	Metric	Value (%)
MLP	Accuracy	87.25
	Sensitivity	79.66
	Specificity	90.93
	Positive predictive value	81.03
	Negative predictive value	90.19
	Accuracy	87.80
Neural Network	Sensitivity	79.10
	Specificity	92.03
	Positive predictive value	82.84
	Negative predictive value	90.05
Deep Learning	Accuracy	89.09
	Sensitivity	81.92
	Specificity	92.58
	Positive predictive value	84.30
	Negative predictive value	91.33

#### **4. Discussion**

Polycystic ovary syndrome is a hormonal disease with short and long-term risks starting from the peripubertal period. In the short term, there are cycle disorders, hirsutism, alopecia, infertility, and pregnancy loss. In the long term, there are risks of type II diabetes, psychosocial problems, cardiovascular diseases, and endometrial cancer. Depression, decreased sexual desire, anxiety, eating disorders and suicide attempts are more common in women with PCOS (Kilic et al., 2020). Diagnosis of PCOS, which occurs in women of reproductive age, is difficult due to gynecological, clinical, and metabolic parameters. In addition to the time required for various clinical tests and ovarian screening to diagnose the disease, financial expenses have also become a challenge for patients with PCOS. As a result of such reasons, women suffer from complications caused by PCOS by neglecting the symptoms in the early stages of the disease. In addition, it is not possible for everyone to afford these tests and scans financially. (Deshmukh et al., 2018). To address this issue, this article is also aimed at identifying risk factors for PCOS development and predicting the disease with the highest accuracy.

ANN is actively used in many applications such as detecting previously unnoticed patterns in medical research data, classifying, controlling medical devices, and determining the characteristics of medical images. The adequacy of artificial intelligence methods has been researched in almost every field of medicine and has the potential to be applied (Gönül et al., 2015).

In recent years, interest in the application of neural networks for problems that cannot be solved with classical techniques has increased and many different ANN models are successfully used in many medical applications. In comparison to existing methods, the MLP method, which is one of these methods, produces very fast results after training. The MLP approach is a non-parametric artificial neural network technique that performs multiple detections and estimation processes. MLP is often preferred for the solution of nonlinear problems (Abdar et al., 2018).

Deep learning methods are an artificial neural network model that is increasingly used today. The Deep learning approach has grown in popularity as graphics processing units (GPU) have improved (Şeker et al., 2017). With the contribution of hardware features that develop over time (especially the participation of graphics processors in the calculation), the deep learning method is used for processing large-sized data, motion detection, face recognition, health technologies (Schmidhuber, 2015). In the area of medicine, deep learning systems assist researchers and medical professionals analyze medical data to treat diseases. It also increases physicians' ability to interpret medical images, accurately diagnose illnesses, and prescribe customized medicines (Keleş, 2018).

According to the findings obtained from this study, the best classification result according to the performance

metrics obtained from the artificial neural networks, MLP and deep learning methods used for the PCOS data set used in the study belongs to the deep learning method. According to the results of the calculated performance criteria, the deep learning method gave successful predictive results in the classification of the data set used in the study. The deep learning method, which gives the best results, obtained the most important risk factors for PCOS. The five most important of these risk factors are Follicle No. (L), Avg. F size (L) (mm), Cycle (R/I), Vit D3 (ng/mL), and Hip (inch) variables.

As a result, PCOS was successfully classified in the light of the findings obtained from the study, and clinical findings were tried to be revealed by giving the risk factors associated with PCOS.

#### **Author Contributions**

Percentages of the author(s) contributions is present below. All authors reviewed and approved final version of the manuscript.

%	Z.K.	F.H.Y.	İ.B.Ç.
C	33	33	34
D	33	33	34
S	33	33	34
DCP	33	33	34
DAI	33	33	34
L	33	33	34
W	33	33	34
CR	33	33	34
SR	33	33	34
PM	33	33	34

C= concept, D= design, S= supervision, DCP= data collection and/or processing, DAI= data analysis and/or interpretation, L= literature search, W= writing, CR= critical review, SR= submission and revision, PM= project management

#### **Conflict of Interest**

The authors declared that there is no potential conflict of interest with respect to the research, authorship, and/or publication of this article.

#### **Ethical Approval/Informed Consent**

The study complied with the Helsinki Declaration, which was revised in 2013. Ethics committee approval was not required for this study because of there is no animal or human research.

#### **References**

- Abdar M, Yen NY, Hung J. CS. 2018. Improving the diagnosis of liver disease using multilayer perceptron neural network and boosted decision trees. *J Med Biol Engin*, 38(6): 953-965.
- Aggarwal, C. C. 2018. Neural networks and deep learning. Springer, 10, 978-973.
- Ayşe A, Berberler ME. 2017. Yapay sinir ağları ile tahmin ve sınıflandırma problemlerinin çözümü için arayüz tasarımı. *Acta Infologica*, 1(2): 55-73.

- Azziz R. 2016. New insights into the genetics of polycystic ovary syndrome. *Nature Rev Endocrinol*, 12(2): 74-75.
- Deshmukh H, Papageorgiou M, Kilpatrick E, Atkin S, Sathyapalan T. 2018. Development of a novel risk prediction and risk stratification score for polycystic ovary syndrome. *Clin Endocrinol*, 90(1): 162-169.
- El-Mahelawi JK, Abu-Daqah JU, Abu-Latifa RI, Abu-Nasser BS, Abu-Naser SS. 2020. Tumor classification using artificial neural networks. *Inter J Acad Engin Res*, 4: 11.
- Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, Dean J. 2019. A guide to deep learning in healthcare. *Nature Med*, 25(1): 24-29.
- Gönül Y, Ulu Ş, Bucak A, Bilir A. 2015. Yapay sinir ağları ve klinik araştırmalarda kullanımı. *Genel Tip Derg*, 25(3): 1-10.
- Karasu S, Saraç Z. 2020. Classification of power quality disturbances with hilbert-huang transform, genetic algorithm and artificial intelligence/machine learning methods. *J Polytech*, 23(4): 1219-1229.
- Kashaninejad M, Dehghani A, Kashiri M. 2009. Modeling of wheat soaking using two artificial neural networks (MLP and RBF). *J Food Engin*, 91(4): 602-607.
- Kayaalp K, Süzen A. 2018. Derin öğrenme ve Türkiye'deki uygulamaları. IKSAD International Publishing House, Adiyaman, Türkiye, 1. Baskı, ss. 89.
- Keleş A. 2018. Derin öğrenme ve sağlık alanındaki uygulamaları. *Elect Turkish Stud*, 13(21): 113-127.
- Kılıç D, Güler, T, Alataş, E. 2020. 2018 Uluslararası kanita dayalı polikistik over sendromu değerlendirme ve yönetim rehberi doğrultusunda uzun dönem risklerin yönetimi. Pamukkale Tıp Derg, 13(2): 453-461.
- LeCun Y, Bengio Y, Hinton G. 2015. Deep learning. *Nature* 521 (7553): 436-444.
- Orhan U, Hekim M, Özer M. 2010. Discretization approach to EEG signal classification using Multilayer Perceptron Neural Network model. In: 15th National Biomedical Engineering Meeting, 21-24 April, Antalya, Turkey, pp. 1-3.
- Öztemel E. 2003. Yapay sinir ağları. Papatya Yayıncılık, İstanbul, Türkiye, 4. Baskı, ss. 232.
- Satish CN, Chew X, Khaw KW. 2020. Polycystic Ovarian Syndrome (PCOS) classification and feature selection by machine learning techniques. *Applied Math Comput Intel*, 9: 65-74.
- Schmidhuber J. 2015. Deep learning in neural networks: An overview. *Neural Networks*, 61: 85-117.
- Steegers-Theunissen RPM, Wiegel RE, Jansen PW, Laven JSE, Sinclair KD. 2020. Polycystic ovary syndrome: A brain disorder characterized by eating problems originating during puberty and adolescence. *Inter J Molec Sci*, 21(21): 8211.
- Şeker A, Diri B, Balık HH. 2017. Derin öğrenme yöntemleri ve uygulamaları hakkında bir inceleme. *Gazi Müh Bilim Derg*, 3(3): 47-64.
- Wang X, Zhao Y, Pourpanah F. 2020. Recent advances in deep learning. *Inter J Machine Learn Cybernet*, 11: 747-750.
- Yüce E, Pabuccu R, Keskin M, Arslanca T, Pabuçcu EG. 2020. Polikistik ovary sendromlu ergen ve yetişkin hastalar arasındaki klinik, endokrinolojik ve biyokimyasal farkların değerlendirilmesi. *Turkish J Reprod Med Surgery*, 4(1): 15-23.