






Detection of Risk Factors of Pcos Patients with Local Interpretable Model-Agnostic Explanations (Lime) Method That an Explainable Artificial Intelligence Model

¹İpek Balıkcı Çiçek , ²Zeynep Küçükakçalı , ³Fatma Hilal Yağın 

^{1,2,3}Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey. (e-mail: {ipek.balikci, zeynep.tunc, hilal.yagin}@inonu.edu.tr).

ARTICLE INFO

Received: Aug.,09.2021

Revised: Sep,28.2021

Accepted: Oct.,19.2021

Keywords:

PCOS
Random forest
Explainable Artificial Intelligence
Local Interpretable Model-agnostic
Explanations (LIME)

Corresponding author: İpek Balıkcı

Çiçek

✉ ipek.balikci@inonu.edu.tr

☎ +90 422 3410660/1337

ISSN: 2548-0650

DOI: <https://doi.org/10.52876/jcs.1004847>

ABSTRACT

In this study, it is aimed to extract patient-based explanations of the contribution of important features in the decision-making process (estimation) of the Random forest (RF) model, which is difficult to interpret for PCOS disease risk, with Local Interpretable Model-Agnostic Explanations (LIME).

In this study, the Local Interpretable Model-Agnostic Annotations (LIME) method was applied to the "Polycystic ovary syndrome" dataset to explain the Random Forest (RF) model, which is difficult to interpret for PCOS risk factors estimation. This dataset is available at <https://www.kaggle.com/prasoonkottarathil/polycystic-ovary-syndrome-pcos>.

Accuracy, sensitivity, specificity, positive predictive value, negative predictive value and balanced accuracy obtained from the Random Forest method were 86.03%, 86.32%, 85.37%, 93.18%, 72.92% and 85.84% respectively. According to the obtained results, the observations whose results were obtained, the values of Follicle (No) L. and Follicle (No) R. in different value ranges were positively correlated with the absence of PCOS. For the observations whose absence of PCOS results were obtained, the variables RBS(mg/dl), bmi_y, fsh_lh, TSH (mIU/L), Endometrium (mm) also played a role in obtaining the results. In addition, for the observations whose results were obtained, the values of Follicle No L and Follicle No R in different value ranges were also found to be positively correlated with PCOS. In addition, beta-HCG(mIU/mL), PRG(ng/mL), RBS(mg/dl), bmi_y, Endometrium (mm), fsh_lh variables also played a role in obtaining the results for PCOS.

When the observations obtained from the results are examined, it can be said that the Follicle (No) L. and Follicle (No) R. variables are the most effective variables on the presence or absence of PCOS. For different value ranges of these two variables, the result of PCOS or not varies. Based on this, it can be said that different values of Follicle (No) L. and Follicle (No) R. variables for PCOS status may be effective in determining the disease.

1. INTRODUCTION

TODAY, tremendous progress in technology has shed light on studies in artificial intelligence (AI) and machine learning (ML). ML methods have achieved great success with predictive models in the analysis of structured data sets in a wide variety of fields, including medical sciences [1]. It is extremely important that ML methods are understandable, explainable and interpretable. However, the inability of researchers to interpret the results of complex models in ML becomes a problem. For ML methods, interpretability is

defined as the degree to which the researcher can understand and interpret the prediction of the model created [2].

Despite the increasing use of predictive models in the medical sciences, clinicians still find it difficult to rely on these models for various reasons [3, 4]. First, most predictive models target specific diseases, and understanding these models depends on clinicians' domain knowledge [4-6]. Second, most models developed by data scientists focus on the model's accuracy in predicting the disease of interest, but models rarely explain these predictions [6, 7]. This is especially true for complex machine learning models such as

Random Forest, Support Vector Machines and Neural Networks, which are described as black boxes [8, 9].

Approaches focusing on Explainable Artificial Intelligence (XAI) have been used in the medical sciences for over two decades. Explainable ML methods that focus on local interpretation, which can be based, for example, on k-NN or decision trees; It has been frequently used recently for its interpretability in predictive models of health-related conditions, including many types of cancer, chronic diseases such as Alzheimer's or diabetes, knee osteoarthritis, and mortality rates from a particular disease [10, 11].

Local Interpretable Model-Agnostic Explanations (LIME), one of the XAI methods, is a popular technique for describing the predictions of black box machine learning models [12]. Because LIME is designed to be model agnostic, it can be applied to many different machine learning models. The model created by the method determines which features in the data are more important on a patient basis, making the results of the model more interpretable [13].

Although many predictive models have been developed to predict the risk of Polycystic Ovary Syndrome (PCOS), one of the most common health problems affecting women of reproductive age, there are often lacking frameworks to build confidence in their predictions [14-16]. PCOS is basically caused by increased body mass index, cycle length, elevated hormone levels, acne, hair loss, hirsutism, infertility, etc. Since the disease has different factors and symptoms, it is important to evaluate the clinical data of PCOS patients on a patient basis, in the diagnosis and treatment of the disease [17, 18]. Therefore, XAI methods can be used to interpret the relative contribution of clinical features for a patient with suspected PCOS.

In this study, it is aimed to extract patient-based explanations of the contribution of important features in the decision-making process (estimation) of the Random forest (RF) model, which is difficult to interpret for PCOS disease risk, with Local Interpretable Model-Agnostic Explanations (LIME).

2. MATERIAL AND METHODS

2.1. Dataset

The In this study, the data set named "Polycystic ovary syndrome" was obtained from <https://www.kaggle.com/prasoonkottathil/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. Random Forest

An The random forest approach proposed by Breiman is a machine learning algorithm with many decision trees which combination of Bagging and Random Subspaces methods [19-21]. In recent years, this method has demonstrated its effectiveness in both regression and classification problems, and it is one of the best machine learning algorithms utilized in a variety of domains [22]. The data set is first randomly divided into two parts: training data for learning and validation data for testing the learning level in the RF algorithm. Following that, many decision trees are generated at random using "boot-strap samples" from the data set. Randomly picked predictors at node locations determine the branching of each tree. The RF Final estimate is the average

of all of the tree's results. As a result, for certain weights, each individual tree has an impact on RF estimation. Because this method shows a "black box" feature, each tree isn't evaluated separately. Because of its ability to randomly accept training data from subsets and create trees with random methods, the RF algorithm outperforms other machine learning algorithms. Furthermore, because training is done on randomly selected different sub-datasets using bootstrap sampling, the random forest algorithm maintains the overfitting level [23, 24].

2.3. Local Interpretable Model-agnostic Explanations

Performance Machine learning and artificial intelligence models are good at prediction accuracy, at process efficiency, and at research productivity. However, current machine learning / artificial intelligence models are often weak in explaining the interpretation process and prediction results. This situation becomes an obstacle in understanding the estimation models created. As a result, clinicians or healthcare workers may think that the results or predictions of a model are not sufficiently descriptive. For this reason, explainable artificial intelligence has started to attract attention recently. In short, explainable artificial intelligence is the totality of methods or techniques that aim to make artificial intelligence applications understandable to users. The purpose of explainable artificial intelligence is to make the computational inferences located behind the decisions of artificial intelligence understandable by researchers. The Local Interpretable Model-Agnostic Explanations (LIME) method is frequently used to ensure the explainability of the artificial intelligence-based models created. The purpose of using LIME is to increase the interpretability and explainability of the created models [25, 26].

In the field of health, the use of classification models to diagnose disease largely depends on the ability to interpretation and explanation of the created models by the researcher. Used for this purpose, LIME provides a patient specific explanation for a particular classification. Thus, it allows any complex classifier to be explained more simply in the clinical setting. LIME, can determine how much each variable in the data contributes to each estimate (patient specific) in the model [27]. Using the LIME method, it can be determined which variables affect each estimation in the model to what degree and in which direction, or which variable has more influence on the results of each estimation in the model compared to other variables. As a result, the LIME method provides explainability for each estimate using any classification model [28].

3. DATA ANALYSIS

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.

4. RESULTS

Descriptive statistics of independent variables included in this study are given in Table 1. 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 Hb (g/dl), FSH(mIU/mL), 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 I
DESCRIPTIVE STATISTICS FOR INDEPENDENT VARIABLES.
PCOS (Y/N)

Variables	PCOS (Y/N)		P-value*
	No	Yes	
	Median (min-max)	Median (min-max)	
Hb (g/dl)	11 (8.5-14.8)	11 (9.4-14)	0.025
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
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

*: Mann Whitney U test

The metrics for the classification performance of Random Forest method in the test phase are given in Table 2. Accuracy, sensitivity, specificity, positive predictive value, negative predictive value and balanced accuracy obtained from the Random Forest method were 86.03%, 86.32%, 85.37%, 93.18%, 72.92% and 85.84% respectively.

TABLE II
VALUES FOR THE METRICS OF THE CLASSIFICATION PERFORMANCE OF RANDOM FOREST METHOD SYSTEMS PARAMETERS.

Method	Metric	Value (%)
Random Forest	Accuracy	86.03
	Sensitivity	86.32
	Specificity	85.37
	Positive predictive value	93.18
	Negative predictive value	72.92
	Balanced Accuracy	85.84

In Figure 1, the findings of the LIME model are given below.

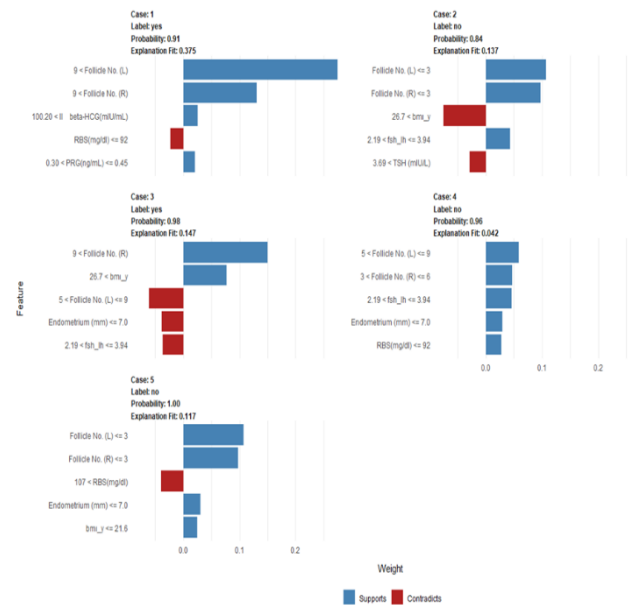


Fig. 1. The findings of the LIME model

For the first patient;
 9 <Follicle No. (L) values are positively associated with PCOS status. Likewise, 9<Follicle No. (R) for the values, there is also a positive correlation with PCOS status. 100.20 <II beta-HCG (mIU / mL) values are positively associated with PCOS status. Finally, 30 <PRG (ng / mL)≤ 45 values are also positively associated with PCOS status. However, values with RBS (mg / dl) ≤0.45 are negatively associated with PCOS status. For this patient, 9 <Follicle No. (L) variable is the most important variable in terms of having PCOS.

For the second patient;
 Follicle No. (L)≤3, Follicle No. (R)≤3 and 2.19<Fsh/Lh≤3.94 values are positively associated with non-PCOS. However, 26.7<BMI and 2.19<Fsh/Lh≤3.94 values are negatively associated with non-PCOS. For this patient, Follicle No. (L)≤3 and Follicle No. (R)≤3 are the most important variables in terms of not having PCOS.

For the third patient;

$9 < \text{Follicle No. (R)}$, $26.7 < \text{BMI}$ values are positively associated with PCOS status. However, $5 < \text{Follicle No. (L)} \leq 9$, $\text{Endometrium (mm)} \leq 7$ and $2.19 < \text{Fsh/Lh} \leq 3.94$ values are negatively associated with PCOS status. For this patient, $9 < \text{Follicle No. (R)}$ variable is the most important variable in terms of having PCOS.

For the fourth patient,

$5 < \text{Follicle No. (L)} \leq 9$, $3 < \text{Follicle No. (L)} \leq 6$, $2.19 < \text{Fsh/Lh} \leq 3.94$, $\text{Endometrium (mm)} \leq 7$ and $\text{RBS (mg / dl)} \leq 92$ values are positively associated with non-PCOS status. For this patient, $5 < \text{Follicle No. (L)} \leq 9$ variable is the most important variable in terms of not having PCOS.

For the fifth patient,

$\text{Follicle No. (L)} \leq 3$, $\text{Follicle No. (R)} \leq 3$, $\text{Endometrium (mm)} \leq 7$ and $\text{BMI} \leq 21.6$ values are positively associated with non-PCOS. $107 < \text{RBS (mg/dl)}$ values are negatively associated with non-PCOS status. For this patient, $\text{Follicle No. (L)} \leq 3$, $\text{Follicle No. (R)} \leq 3$ variables are the most important variables in terms of not having PCOS.

5. DISCUSSION

The most frequent endocrine disorder in women of reproductive age is polycystic ovarian syndrome (PCOS). PCOS affects roughly 6–10% of women of reproductive age, depending on the diagnostic criteria used. A combination of clinical indications of menstrual abnormalities or anovulation, clinical or biochemical hyperandrogenism, and polycystic ovaries is used to diagnose PCOS. It is often diagnosed in the reproductive years of life when women with PCOS are meted with infertility, or because of hyperandrogenism symptoms such as acne, alopecia androgenica, and hirsutism [29]. Because the disease is complex and multifactorial, diagnosing it might be challenging owing to overlapping symptoms. Multiple etiological factors have been implicated in PCOS. Despite these in the care and diagnosis of PCOS, progress has been made [30].

Artificial Intelligence (AI) has gained significant traction in recent years, and if properly handled, it has the potential to exceed expectations across a wide range of application industries [31]. Current artificial intelligence technologies, on the other hand, are often poor at describing the interpretation process and predicting outcomes. This condition makes it difficult to comprehend the estimation models that have been developed. As a result, explainable artificial intelligence has lately begun to gain traction. The goal of explainable artificial intelligence is to make the computational conclusions that underpin AI decisions understandable to ordinary users and academics [32].

LIME (Local Interpretable Model-Agnostic Explanations) is a common methodology for making black box Machine Learning (ML) algorithms more interpretable and explainable. LIME often generates an explanation for a single prediction made by any ML model by learning a simpler interpretable model (e.g. linear classifier) around the prediction by randomly perturbing simulated data around the instance and obtaining feature importance through feature selection. LIME and similar local algorithms have gained popularity due to their simplicity. The LIME approach can be used to discover which variables affect each estimation in the model to what extent and in which direction, as well as which variable has a greater impact on the model's outcomes than other factors. This gives a detailed explanation for each observation, allowing any complex classifier to be explained in a straightforward manner [28].

According to the interpreted results, the observations whose results were obtained, the values of Follicle (No) L. and Follicle (No) R. in different value ranges were positively correlated with the absence of PCOS. In addition, these two variables are the most important variables for the absence of PCOS for our observations. For the observations whose results were obtained, the variables RBS(mg/dl), bmi_y, fsh_lh, TSH (mIU/L), Endometrium (mm) also played a role in obtaining the results.

In addition, for the observations whose results were obtained, the values of Follicle No L and Follicle No R in different value ranges were also found to be positively correlated with PCOS. Likewise, the most important variables for PCOS status are Follicle No. (L) and Follicle No. (R) variables. In addition, beta-HCG(mIU/mL), PRG(ng/mL), RBS(mg/dl), bmi_y, Endometrium (mm), fsh_lh variables also played a role in obtaining the results for PCOS.

When the observations obtained from the results are examined, it can be said that the Follicle (No) L. and Follicle (No) R. variables are the most effective variables on the presence or absence of PCOS. For different value ranges of these two variables, the result of PCOS or not varies. Based on this, it can be said that different values of Follicle (No) L. and Follicle (No) R. variables for PCOS status may be effective in determining the disease.

According to the results, the values of Follicle No L and Follicle No R variables tend to increase in the case of PCOS.

REFERENCES

- [1] Yu, K. H., Beam, A. L., & Kohane, I. S. (2018). Artificial intelligence in healthcare. *Nature biomedical engineering*, 2(10), 719-731.
- [2] Sokol, K., Santos-rodriguez, R., Hepburn, A., & Flach, P. (2019). *Surrogate Prediction Explanations Beyond LIME*. no. HCML.
- [3] Kononenko, I. (2001). Machine learning for medical diagnosis: history, state of the art and perspective. *Artificial Intelligence in medicine*, 23(1), 89-109.
- [4] Deo, R. C. (2015). Machine learning in medicine. *Circulation*, 132(20), 1920-1930.
- [5] He, D., Mathews, S. C., Kallou, A. N., & Hutfless, S. (2014). Mining high-dimensional administrative claims data to predict early hospital readmissions. *Journal of the American Medical Informatics Association*, 21(2), 272-279.
- [6] Pederson, J. L., Majumdar, S. R., Forhan, M., Johnson, J. A., McAlister, F. A., & PROACTIVE Investigators. (2016). Current depressive symptoms but not history of depression predict hospital readmission or death after discharge from medical wards: a multisite prospective cohort study. *General hospital psychiatry*, 39, 80-85.
- [7] Futoma, J., Morris, J., & Lucas, J. (2015). A comparison of models for predicting early hospital readmissions. *Journal of biomedical informatics*, 56, 229-238.
- [8] Katuwal, G. J., & Chen, R. (2016). Machine learning model interpretability for precision medicine. *arXiv preprint arXiv:1610.09045*.
- [9] Bastani, O., Kim, C., & Bastani, H. (2017). Interpreting blackbox models via model extraction. *arXiv preprint arXiv:1705.08504*.
- [10] Stiglic, G., Kocbek, P., Fijacko, N., Zitnik, M., Verbert, K., & Cilar, L. (2020). Interpretability of machine learning-based prediction models in healthcare. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(5), e1379.
- [11] Escalante, H. J., Escalera, S., Guyon, I., Baró, X., Güçlütürk, Y., Güçlü, U., ... & van Lier, R. (Eds.). (2018). *Explainable and interpretable models in computer vision and machine learning*. Cham: Springer International Publishing.
- [12] Garreau, D., & Luxburg, U. (2020, June). Explaining the explainer: A first theoretical analysis of LIME. In *International Conference on Artificial Intelligence and Statistics* (pp. 1287-1296). PMLR.
- [13] Hu, L., Chen, J., Nair, V. N., & Sudjianto, A. (2018). Locally interpretable models and effects based on supervised partitioning (LIME-SUP). *arXiv preprint arXiv:1806.00663*.
- [14] Mehrotra, P., Chatterjee, J., Chakraborty, C., Ghoshdastidar, B., & Ghoshdastidar, S. (2011, December). Automated screening of

- polycystic ovary syndrome using machine learning techniques. In 2011 Annual IEEE India Conference (pp. 1-5). IEEE.
- [15] Denny, A., Raj, A., Ashok, A., Ram, C. M., & George, R. (2019, October). I-HOPE: detection and prediction system for polycystic ovary syndrome (PCOS) using machine learning techniques. In TENCON 2019-2019 IEEE Region 10 Conference (TENCON) (pp. 673-678). IEEE.
- [16] Meena, K., Manimekalai, M., & Rethinavalli, S. (2015). Correlation of Artificial Neural Network Classification and NFRS Attribute Filtering Algorithm for PCOS Data. *Int. J. Res. Eng. Technol.*, 4(3), 519-524.
- [17] Vikas, B., Anuhyu, B. S., Chilla, M., & Sarangi, S. (2018). A Critical Study of Polycystic Ovarian Syndrome (PCOS) Classification Techniques. *International Journal of Computational Engineering & Management*, 21(4), 1-7.
- [18] Khasar-Miller, M. D., Nixon, C., Boots, L. R., Go, R. C., & Azziz, R. (2001). Prevalence of polycystic ovary syndrome (PCOS) in first-degree relatives of patients with PCOS. *Fertility and sterility*, 75(1), 53-58.
- [19] Breiman, L. (1996). Bagging predictors. *Machine learning*, 24(2), 123-140.
- [20] Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.
- [21] Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE transactions on pattern analysis and machine intelligence*, 20(8), 832-844.
- [22] Izquierdo-Verdiguier, E., & Zurita-Milla, R. (2020). An evaluation of Guided Regularized Random Forest for classification and regression tasks in remote sensing. *International Journal of Applied Earth Observation and Geoinformation*, 88, 102051.
- [23] Prasad, A. M., Iverson, L. R., & Liaw, A. (2006). Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems*, 9(2), 181-199.
- [24] Panov, P., & Džeroski, S. (2007, September). Combining bagging and random subspaces to create better ensembles. In *International Symposium on Intelligent Data Analysis* (pp. 118-129). Springer, Berlin, Heidelberg.
- [25] Shi, S., Zhang, X., & Fan, W. (2020). A modified perturbed sampling method for local interpretable model-agnostic explanation. *arXiv preprint arXiv:2002.07434*.
- [26] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016, August). "Why should i trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1135-1144).
- [27] Kumarakulasinghe, N. B., Blomberg, T., Liu, J., Leao, A. S., & Papapetrou, P. (2020, July). Evaluating local interpretable model-agnostic explanations on clinical machine learning classification models. In *2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)* (pp. 7-12). IEEE.
- [28] Zafar, M. R., & Khan, N. M. (2019). DLIME: A deterministic local interpretable model-agnostic explanations approach for computer-aided diagnosis systems. *arXiv preprint arXiv:1906.10263*.
- [29] McLuskie, I., & Newth, A. (2017). New diagnosis of polycystic ovary syndrome. *BMJ: British Medical Journal*, 356.
- [30] Khan, M. J., Ullah, A., & Basit, S. (2019). Genetic basis of polycystic ovary syndrome (PCOS): current perspectives. *The application of clinical genetics*, 12, 249.
- [31] Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82-115.
- [32] Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G. Z. (2019). XAI—Explainable artificial intelligence. *Science Robotics*, 4(37).

BIOGRAPHIES

İpek Balıkcı Çiçek obtained her BSc. degree in mathematics from Çukurova University in 2010. She received MSc. degree in biostatistics and medical informatics from the Inonu University in 2018. She currently continues Ph.D. degrees in biostatistics and medical informatics from the Inonu University. In 2014, she joined the Department of Biostatistics and Medical Informatics at Inonu University as a researcher assistant. Her research interests are cognitive systems, data mining, machine learning, deep learning.

Zeynep Küçükakçalı obtained her BSc. degree in mathematics from Çukurova University in 2010. She received MSc. degree in biostatistics and medical informatics from the Inonu University in 2018. She currently continues Ph.D. degrees in biostatistics and medical informatics from the Inonu University. In 2014, she joined the Department of Biostatistics and Medical Informatics at Inonu

University as a researcher assistant. Her research interests are cognitive systems, data mining, machine learning, deep learning.

Fatma Hilal Yağın obtained her BSc. degree in Statistics from Gazi University in 2017. She received MSc. degree in biostatistics and medical informatics from the Inonu University in 2020. She currently continues Ph.D. education in biostatistics and medical informatics from the Inonu University. In 2019, she joined the Department of Biostatistics and Medical Informatics at Inonu University as a researcher assistant. Her research interests are cognitive systems, data mining, machine learning, deep learning, and image processing