Multilayer stacking for polycystic ovary syndrome diagnosis

Kazi Abu Taher[1], Samia Ahmed[1], Jannatul Ferdous Esha[1], Md. Sazzadur Rahman[2], A. S. M. Sanwar Hosen[3]

[1]Department of Information and Communication Technology, Bangladesh University of Professionals, Dhaka, Bangladesh
[2]Institute of Information Technology, Jahangirnagar University, Savar, Bangladesh
[3]Department of Artificial Intelligence and Big Data, Woosong University, Daejeon, South Korea

Article Info

Article history:

Received May 4, 2024
Revised Nov 29, 2024
Accepted Jan 27, 2025

Keywords:

Hormonal disease
Machine learning
PCOS
Stacking
Women health

ABSTRACT

Polycystic ovary syndrome (PCOS) is a complicated hormonal condition that is experienced by women. Despite extensive research, the precise reason behind PCOS remains unknown, and effective treatments are still lacking. Thus, early diagnosis and treatment have a significant positive impact on the health of women. Recently, there has been remarkable performance demonstrated by machine learning (ML)-based detection models for PCOS identification. They are fast and low cost compared to the traditional processes. In this work, a multi stacking PCOS detection model is proposed using K-fold cross validation. The model uses three different ML algorithms namely: naïve Bayes (NB), random forest (RF), and logistic regression (LR) as base classifiers and a neural network, multi-layer perception (MLP) as meta model. This approach utilizes two feature selection techniques and compares the performances on the stacking methods. Among the two feature selection techniques, Pearson correlation approach performed better with average 98.79% accuracy, 99.17% sensitivity, 98.40% specificity, and 98.79% f1-score.

Corresponding Author:

A. S. M. Sanwar Hosen
Department of Artificial Intelligence and Big Data, Woosong University
Daejeon 34606, South Korea
Email: sanwar@wsu.ac.kr

1. INTRODUCTION

Polycystic ovary syndrome (PCOS) ranks among the prevalent hormonal imbalances found in women. According to World Health Organization (WHO), 8-13% of women of reproductive age suffers from PCOS and approximately 13% of them remain undetected [1]. Male hormone levels are greater in PCOS-affected women than in healthy individuals, which may lead to infertility. PCOS also causes cardiovascular disease, type-2 diabetes, sleep apnea, and trimester miscarriage [2], [3]. The primary signs of PCOS consist of irregular menstrual cycle, immature ovarian eggs, various cysts on the ovaries, difficulty in conceiving, an excess development of hair often on the back, buttocks, chest, and upper lip, weight gain, hair fall, oily or acne-prone skin, obesity, skin pigmentation, and skin discoloration [4]–[6]. Manual detection of PCOS includes hormonal tests, symptom aggregation, and detection of cysts in ultrasound images. However, these traditional approaches are expensive, lengthy, and susceptible to human error. Therefore, a shift towards machine learning (ML)-based methods can solve the problems. Recently, several articles have been published to diagnosis PCOS using ML approaches. Based on patient clinical data, the research in [7], [8] used techniques including support vector machine (SVM), classification and regression trees (CART), naïve Bayes (NB), random forest (RF), and logistic regression (LR). Denny *et al.* [8] also used k-nearest neighbor (KNN) to detect. RF algorithm has the maximum accuracy of 96% [7]. Prapty and Shitu [9] suggested a process that makes use of a few ML techniques, including NB, KNN, RF, and SVM. Although RF provided the most accuracy, NB and RF delivered comparable performance. Mehrotra *et al.* [10] proposed a contemporary method of PCOS screening. Two classifiers such as: LR and Bayesian classifier are used. The Bayesian classifier outperforms LR. SVM, NB, KNN, and ensemble approaches were employed in [11] for PCOS classification. Ensemble techniques consist of SVM and KNN. To identify if a woman has PCOS or not. Thakre *et al.* [12] employed five different ML classifiers: RF, SVM, LR, NB, and KNN. RF surpassed with an accuracy rate of 90.9%. Ten ML methods were assessed by Panda *et al.* [13]. These models included NB, KNN, stochastic gradient descent (SGD), multi-layer perception (MLP), RF, LR, NB, SVM, linear regression, and Bayesian Ridge. RF performed best with an accuracy of 92%. The efficacy of methods like RF, convolutional neural network (CNN), and SVM was demonstrated in [14] as they presented ML strategies for PCOS identification.

Another approach of ML is boosting which is applied to several models. A method using hybrid random forest logistic regression (HRFLR), extreme gradient boosting with random forest (XGBRF), linear support vector machine (LSVM), light gradient boosting model (LGB), and CatBoost model was proposed in [15]. Bhat [16] put out a novel method that combines CatBoost and XGBRF models. Different classifiers, including gradient boost (GB), RF, LR, HRFLR, MLP, SVM, and decision tree (DT) were employed as base methods to assess the outcomes. With an emphasis on the XGBoost algorithm, Avasthi *et al.* [17] explored the accurate identification of PCOS using ML. With a testing accuracy of 96%, the XGBoost algorithm demonstrated impressive effectiveness. Modi and Kumar [18] applied several ML models such as SVM, RF, DT, NB, LR, GB, Catboost, and adaptive boosting (Adaboost). Among all the methods, CatBoost performed the best. Explainable artificial intelligence (XAI) is a very important concept to clarify ML algorithms' decision-making process. XAI was also utilized in [19]. Utilizing NB, KNN, SVM with RBF and linear kernel, dense neural network (DNN), and RF. Stacking is another noteworthy ML technique. Suha and Islam [20] presented a stacking model that uses one bagging or boosting ensemble ML model as the stacked model's meta-learner and five regular ML models as base learners. Kumari *et al.* [21] employed six classifiers to implement SmS hybrid models. They utilized LR, DT, RF, SVM, NB, and AdaBoost as base learners at the base level. At the meta-level, each of these classifiers was taken into consideration independently. Using fourteen ML techniques, Akhtar *et al.* [22] compared and determined which model was the best. Using synthetic minority over-sampling technique (SMOTE), data was re-sampled. Classifiers such as GB, AdaBoost, CatBoost, RF, KNN, SVM, LR, NB, voting, and methods have been used to compare the results. The research in [23], [24] used AdaBoost, KNN, LR, RF, DT, NB, SVM, and XGboost to obtain the best model.

In this work, a data-driven strategy is proposed to identify PCOS using a stacking based model. K-fold cross validation is used to train and validate the model following preprocessing and feature selection. There are one meta layer and two base layers in the model. Different ML-based models are employed in the base layers, whereas MLP serves as the meta layer. The principal findings of this study are given as follows:

- A novel multilayer stacking approach is proposed, combining various classification algorithms at different base layers with an MLP at the meta-layer for PCOS detection.
- The experiments are conducted on a new dataset, which has not been utilized previously.
- The impact of two different feature selection techniques is compared.
- This work individually trains the base layer algorithms using this dataset and compares the performance with the proposed model.

This paper is structured as follows: section 2 of the paper covers the research method. It explains the multilayer stacking paradigm that has been suggested for PCOS detection. In section 3, the experimental data are examined and discussed. Conclusions and future scope are drawn in section 4.

2. RESEARCH METHOD

This experiment was carried out using Python 3.8 and cloud-based Google Colab, employing libraries such as scikit-learn, Keras, TensorFlow, NumPy, Pandas, and imbalanced-learn. ML model training was sped up with a GPU (NVIDIA Tesla T4). This work outlines a structure that consists of data collection, data preprocessing, feature selection, training proposed model, and evaluation of the result. Figure 1 visualizes the schematic diagram of this work. In Figure 1(a) the flowchart of this work is visualized. It starts with data collection. After data collection, data is preprocessed using various approaches. Then important features are selected. Next, the proposed model is trained and finally, the result is evaluated using different performance parameters. In Figure 1(b) the architecture of this work is presented. The simulation and experimental setup can be found on the GitHub [25].

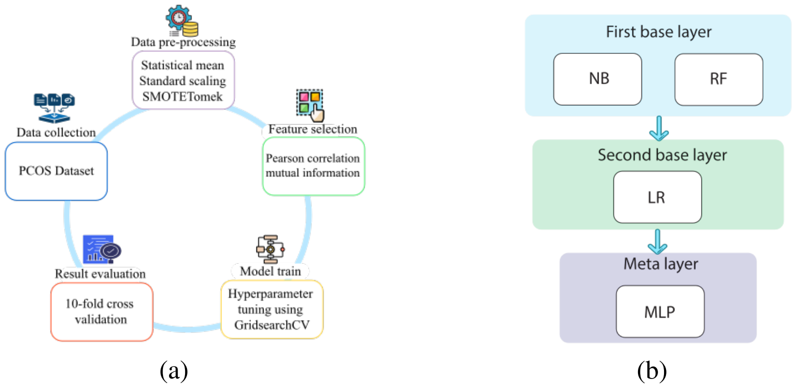

Figure 1. Diagrammatic representation of the suggested work: (a) the flowchart of this work for diagnosis of PCOS and (b) architecture of the proposed model for PCOS detection using multilayer stacking

2.1. Data collection

The dataset was gathered using Kaggle. It is a freely accessible portal with a range of datasets. There are 44 feature columns and 2000 rows in the gathered PCOS dataset [26]. In this dataset, the positive and negative classess include 608 and 1392 instances respectively. The majority class, 69.6% of the data shows the percentage of women not having PCOS, while 30.4% of the minority class's women with PCOS.

2.2. Data preprocessing

It ensures data quality, reliability, reduces overfitting, and improves model performance. In this work, firstly the missing values are checked. Not-a-number (NaN) or null values are frequently used in datasets to indicate missing values. It was observed that there were some NaN values in the dataset. A statistical metrics named mean was used to impute the missing values. The features were removed that were not significant enough such as 'SI/No' and 'Patient file no'. Next, data scaling was carried out. The data underwent standard scaling prior to further analysis. In standard scaling, the values of each feature are centered around a standard deviation of one with a mean of zero [27]. After scaling, the dataset is balanced. The significance of balancing datasets is underscored by the possibility of biased outcomes arising from dataset imbalances. In this work, SMOTETomek is used. It is a resampling technique that uses under sampling with Tomek link and oversampling with SMOTE [28]. After resampling, PCOS and non-PCOS classes included 1295 instances.

2.3. Feature selection

There are a lot of unnecessary, redundant, and noisy features in real-world data. The curse of dimensionality is a significant problem that arises when data mining (DM) and ML techniques are used on high-dimensional data. Data gets sparser in high-dimensional space. Therefore, learning models tend to overfit when they have a lot of features, which can lead to a drop in performance on test data. Removing such features through feature selection saves computing and storage costs without causing a large loss of data or a decrease in learning efficiency. In this work, two different feature selection techniques, Pearson correlation [29] and mutual information [30] methods applied to selecting important features to make the model robust. For Pearson correlation, the threshold value was set to 80. The features that are correlated more than 80% are considered highly correlated and to prevent repetition in the dataset, one of the correlated features is eliminated. On the other hand, the top 20 features were selected in mutual information method. The features were selected based on mutual information scores.

2.4. Proposed model

A novel multi-layer stacking model is presented based on stacking principles [31]. The proposed approach combines diverse ML algorithms in two distinct base layers, with a neural network functioning as the meta layer. In the first base layer, the NB [32] and RF [33] classifiers are trained separately. Although NB is sensitive to imbalanced data, it is a straightforward and computationally cheap technique. SMOTETomek is used to address that problem. RF can decrease overfitting and manage complex models with a high success rate. The first layer is balanced by these two algorithms. These algorithms then produce predictions, which are saved and used as training data for the second base layer model. Due to its high efficiency, LR [34] is used in the second base layer. A meta model that is MLP [35] is trained based on the second base layer's predictions. The ultimate result is made up of the results that the meta model provides. GridSearchCV was used for the meta layer hyperparameter tuning. It is a scikit-learn function that essentially considers all possible combinations of the candidates to find the optimal set of hyperparameters by training. To optimize the model, the Adam optimizer was used at the meta layer. To make the model more reliable and assess the performance, K-fold [36] cross validation is applied. It also maximizes the utilization of available data. In this manuscript, the value of K is considered 10 that means the dataset is split up into 10 folds. Every fold serves as a validation set once, and the training set is made up of the remaining folds.

3. RESULTS AND DISCUSSION

This section presents the experimental result of the proposed model along with a comparison of results obtained from other existing models. Accuracy, sensitivity, specificity, f1-score, and area under the curve (AUC) are used to assess the model. The suggested model is contrasted with two different feature selection methods. In case of Pearson correlation method, after ten fold cross validation the model achieved average 98.79% accuracy, 99.17% sensitivity, 98.40% specificity, 98.79% f1-score, and 0.9879 AUC. Tables 1 and 2 exhibits varying levels of performance parameters including accuracy, sensitivity, specificity, f1-score and AUC across different folds for Pearson correlation and mutual information respectively.

Table 1. The performance of the model in each fold using Pearson correlation feature selection method

Fold	Accuracy	Sensitivity	Specificity	F1-score	AUC
1	0.9871	0.9829	0.9914	0.9871	0.9871
2	0.9924	1.0	0.9948	0.9924	0.9924
3	0.9921	1.0	0.9843	0.9922	0.9921
4	0.9964	1.0	0.9929	0.9964	0.9964
5	0.9860	1.0	0.9720	0.9862	0.9860
6	0.9885	0.9923	0.9847	0.9885	0.9885
7	0.9860	0.9860	0.9860	0.9860	0.9860
8	0.9767	0.9922	0.9612	0.9770	0.9767
9	0.9875	0.9916	0.9833	0.9875	0.9874
10	0.9863	0.9727	1.0	0.9862	0.9863
Average	0.9879	0.9917	0.9840	0.9879	0.9879

Table 2. The performance of the model in each fold using mutual information feature selection method

Fold	Accuracy	Sensitivity	Specificity	F1-score	AUC
1	0.9572	0.9145	1.0	0.9553	0.9572
2	0.9772	1.0	0.9545	0.9777	0.9772
3	0.9726	0.9921	0.9531	0.9531	0.9726
4	0.9751	0.9787	0.9716	0.9752	0.9751
5	0.9860	0.9720	1.0	0.9858	0.9860
6	0.9656	0.9770	0.9541	0.9660	0.9656
7	0.9755	0.9650	0.9860	0.9752	0.9755
8	0.9534	1.0	0.9069	0.9555	0.9534
9	0.9625	0.9416	0.9833	0.9617	0.9624
10	0.9761	0.9659	0.9863	0.9759	0.9761
Average	0.9701	0.9701	0.9696	0.9701	0.9701

The model achieved 97.01%, 97.01%, 96.96%, 97.01%, and 0.9701 of accuracy, sensitivity, specificity, f1-score, and AUC respectively in average using mutual information feature selection approach. It is visible that Pearson correlation approach has performed better. Accuracy simply measures the overall correctness of predictions made by model. Sensitivity represents percentage of true positives cases that were correctly identified by model. Specificity refers to the correctly identified percentage of true negatives. In proposed approach, apart from accuracy, value of sensitivity and specificity is also high. For Pearson correlation average sensitivity is 99.17%. That means in 99.17% cases model identified PCOS affected data correctly. Using mutual information, the result decreased by about 2.2%. In the case of specificity, Pearson correlation correctly identified 98.40% non PCOS cases as true negatives which is 1.46% higher than mutual information.

In the investigation, the training loss is calculated in every fold to assess the proposed model. A comparison-based illustration of the model's training loss for the two feature selection techniques is shown in Figure 2. While both approaches yield good results, Pearson correlation outperforms mutual information by a little margin. Initially, the loss for both approaches was 0.7. The loss considerably reduced in the first few folds. Following the fourth fold, the loss converges close to zero and is linear in each fold. The training loss is likewise reasonable after employing mutual information. It is marginally higher in each fold when compared to Pearson correlation. Following ten folds, the training loss approaches 0.1. To evaluate the performance, The proposed model is contrasted with the base classifiers independently. From Figure 3, the result of base classifiers on the same dataset is found. Among NB, RF, and LR classifiers, LR performed comparatively better with 95.8%, 94.5%, 97.1%, and 95.7% of accuracy, sensitivity, specificity, and f1-score respectively. However, the proposed model performed best in comparison with other classifiers.

Table 3 compares the performance of the previous studies with the proposed multi layer stacking approach. Evidently, the proposed work has the best performance compared to previous studies. After using multilayer stacking, the proposed model achieved 98.7% accuracy, 99.17% sensitivity, 98.40% specificity, and 98.79% f1-score.

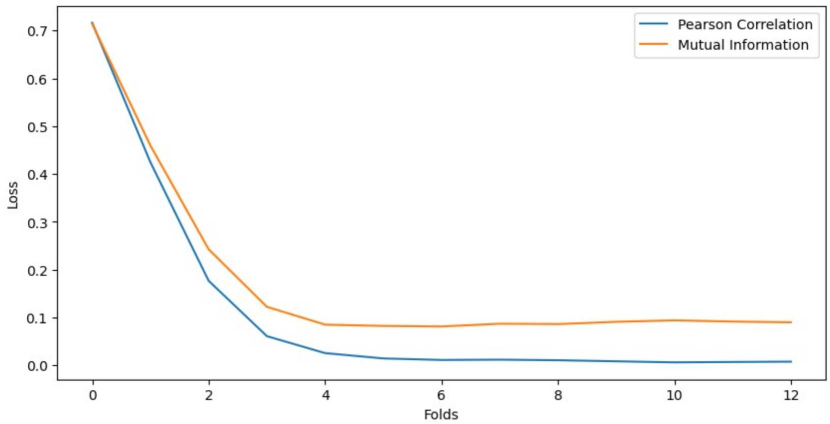

Figure 2. Training loss of the proposed model in each fold for two different feature selection methods

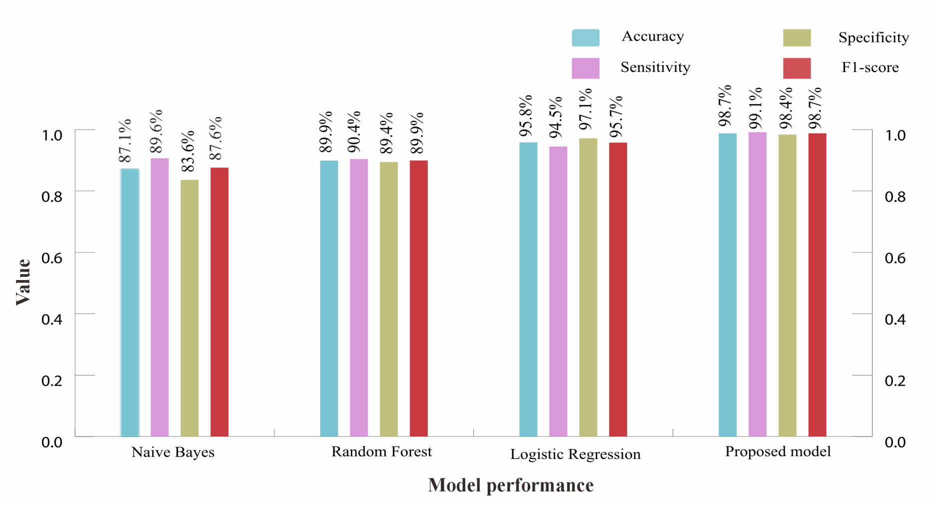

Figure 3. Comparison of performance of the proposed model with the base classifiers

Table 3. Comparison of the proposed model's performance with previous studies

References	Model	Accuracy (%)	Sensitivity/Recall (%)	Specificity (%)	Precision (%)	F1-score (%)
[7]	RF	96	95	-	96	96
[8]	RF	89.02	74.19	98.03	95.83	41.82
[9]	RF	93.5	84	-	84	84
[10]	Bayesian classifier	93.93	92.85	94.23	-	-
[11]	Ensemble (SVM+KNN)	97.22	-	-	98.70	98.06
[12]	RF	90.9	97	-	89.1	92
[15]	CatBoost	92	84	-	95	89
[16]	CatBoost	95	90	-	83	86
[18]	CatBoost	93	96	-	95	96
[20]	Stacking	95.7	95.7	-	95.6	96
[21]	Stacking	90.24	89.93	-	90	89.82
[23]	Stacking	98	98	-	97	98
[24]	Stacking	98.87	98.87	-	98	98.89
Proposed approach	Multi-layer stacking	98.7	99.17	98.40	-	98.79

4. CONCLUSION

This model incorporates three-layer ensemble approach and demonstrated superior performance, especially with Pearson correlation feature selection technique. The proposed model acquired average test accuracy of 98.7%. The training loss curves showed stable convergence, highlighting model's robust learning. The models' comparisons explain how well the proposed model predicts. Broadly, this study emphasizes potential of ensemble learning and feature selection in developing accurate PCOS prediction models with clinical implications for early diagnosis and management. In future, different base classifiers and meta models can be adopted. The effectiveness of model may also be evaluated using explainable AI and large dataset.

FUNDING INFORMATION

This work was supported by the Woosong University Academic Research Fund 2024, South Korea.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Kazi Abu Taher	✓	✓	✓	✓	✓	✓		✓		✓		✓		
Samia Ahmed	✓		✓			✓		✓	✓		✓			
Jannatul Ferdous Esha	✓		✓			✓		✓	✓		✓			
Md. Sazzadur Rahman		✓		✓	✓					✓		✓	✓	
A. S. M. Sanwar Hosen		✓			✓		✓			✓		✓	✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal Analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject Administration
Fu : **Fu**nding Acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The supporting data of this study are openly available in [Kaggle] at https://www.kaggle.com/datasets/cm037divya/pcos-dataset, reference number [26].

REFERENCES

[1] WHO, "Polycystic ovary syndrome," *World Health Organization*, 2025. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/polycystic-ovary-syndrome

[2] S. Ahmed *et al.*, "A review on the detection techniques of polycystic ovary syndrome using machine learning," *IEEE Access*, vol. 11, pp. 86522–86543, 2023, doi: 10.1109/ACCESS.2023.3304536.

[3] A. Chaudhuri, "Polycystic ovary syndrome: Causes, symptoms, pathophysiology, and remedies," *Obesity Medicine*, vol. 39, 2023, doi: 10.1016/j.obmed.2023.100480.

[4] S. Yeruva, I. Gurrala, R. S. Myakala, N. Agarwal, S. Rapolu, and J. Ding, "KNOw PCOS," in *Third International Conference on Advances in Computer Engineering and Communication Systems*, 2023, pp. 533–546, doi: 10.1007/978-981-19-9228-5_45.

[5] S. Chen, N. Tai, C. Fan, J. Liu, and S. Hong, "Sequence-component-based current differential protection for transmission lines connected with IIGs," *IET Generation, Transmission and Distribution*, vol. 12, no. 12, 2018, doi: 10.1049/IET-GTD.2017.1507.

[6] S. Parhizi, H. Lotfi, A. Khodaei, and S. Bahramirad, "State of the art in research on microgrids: A review," *IEEE Access*, vol. 3, pp. 890–925, 2015, doi: 10.1109/ACCESS.2015.2443119.

[7] M. M. Hassan and T. Mirza, "Comparative analysis of machine learning algorithms in diagnosis of polycystic ovarian syndrome," *International Journal of Computer Applications*, vol. 175, no. 17, pp. 42–53, 2020, doi: 10.5120/ijca2020920688.

[8] A. Denny, A. Raj, A. Ashok, C. M. Ram, and R. George, "I-HOPE: Detection and prediction system for polycystic ovary syndrome (PCOS) using machine learning techniques," in *2019 IEEE Region 10 Conference*, 2019, doi: 10.1109/TENCON.2019.8929674.

[9] A. S. Prapty and T. T. Shitu, "An efficient decision tree establishment and performance analysis with different machine learning approaches on polycystic ovary syndrome," in *2020 23rd International Conference on Computer and Information Technology (ICCIT)*, 2020, pp. 1–5, doi: 10.1109/ICCIT51783.2020.9392666.

[10] P. Mehrotra, J. Chatterjee, C. Chakraborty, B. Ghoshdastidar, and S. Ghoshdastidar, "Automated screening of polycystic ovary syndrome using machine learning techniques," in *2011 Annual IEEE India Conference*, 2011, doi: 10.1109/INDCON.2011.6139331.

[11] M. Alagarsamy, N. Shanmugam, D. P. Mani, M. Thayumanavan, K. K. Sundari, and K. Suriyan, "Detection of polycystic syndrome in ovary using machine learning algorithm," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 11, no. 1, pp. 246–253, 2023.

[12] V. Thakre, S. Vedpathak, K. Thakre, and S. Sonawai, "PCOcare: PCOS detection and prediction using machine learning algorithms," *Bioscience Biotechnology Research Communications*, vol. 13, no. 14, pp. 240–244, 2020, doi: 10.21786/bbrc/13.14/56.

[13] O. K. Panda *et al.*, "Development and analysis of machine learning models for polycystic ovary syndrome detection," in *2024 1st International Conference on Cognitive, Green and Ubiquitous Computing*, 2024, doi: 10.1109/IC-CGU58078.2024.10530738.

[14] S. Juneja, B. S. Bhati, S. Atwal, M. Anand, and S. Maiti, "Unravelling the enigma of polycystic ovary syndrome (PCOS): using MLAlgorithms," in *2024 2nd International Conference on Disruptive Technologies (ICDT)*, 2024, pp. 1480–1485, doi: 10.1109/ICDT61202.2024.10489567.

[15] S. Alshakrani, S. Hilal, and A. M. Zeki, "Hybrid machine learning algorithms for polycystic ovary syndrome detection," in *2022 International Conference on Data Analytics for Business and Industry*, 2022, doi: 10.1109/ICDABI56818.2022.10041525.

[16] S. A. Bhat, "Detection of polycystic ovary syndrome using machine learning algorithms," *M.Sc. Data Analytics*, School of Computing, National College of Ireland, Dublin, Ireland, 2021.

[17] V. Avasthi, A. Kumar, A. Bhardwaj, and T. Jain, "Empowering women's health: Machine learning for PCOS detection and prediction," in *2024 International Conference on Distributed Computing and Optimization Techniques (ICDCOT)*, 2024, pp. 1–6, doi: 10.1109/ICDCOT61034.2024.10516171.

[18] N. Modi and Y. Kumar, "Detection and classification of polycystic ovary syndrome using machine learning-based approaches," in *2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, 2024, pp. 1–6, doi: 10.1109/IATMSI60426.2024.10503222.

[19] P. Jain, R. K. Mishra, A. Deep, and N. K. Jain, "Xplainable AI for deep learning model on PCOS analysis," in *XAI Based Intelligent Systems for Society 5.0*, 2023, pp. 131–152, doi: 10.1016/B978-0-323-95315-3.00012-7.

[20] S. A. Suha and M. N. Islam, "Exploring the dominant features and data-driven detection of polycystic ovary syndrome through modified stacking ensemble machine learning technique," *Heliyon*, vol. 9, no. 3, 2023, doi: 10.1016/j.heliyon.2023.e14518.

[21] R. Kumari, J. Singh, and A. Gosain, "SmS: SMOTE-stacked hybrid model for diagnosis of polycystic ovary syndrome using feature selection method," *Expert Systems with Applications*, vol. 225, 2023, doi: 10.1016/j.eswa.2023.120102.

[22] M. Akhtar, K. A. Ahmed, and F. Al-Islam, "An improved prediction of polycystic ovary syndrome using SMOTE-based oversampling and stacking classifier," in *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2023, pp. 1–6, doi: 10.1109/ICCCNT56998.2023.10307634.

[23] V. V. Khanna, K. Chadaga, N. Sampathila, S. Prabhu, V. Bhandage, and G. K. Hegde, "A distinctive explainable machine learning framework for detection of polycystic ovary syndrome," *Applied System Innovation*, vol. 6, no. 2, 2023, doi: 10.3390/asi6020032.

[24] K. Saraswat, D. K. Nama, G. Sharma, and G. Lavania, "PCOS detection machine learning model," *Pratibodh: A Journal for Engineering*, vol. 2023, pp. 1–3, 2024.

[25] S. Ahmed, "Multi-layer stacking," *GitHub*, 2025. [Online]. Available: https://github.com/samiaomi/Multi-layer-stacking

[26] Divya, "PCOS dataset," *Kaggle*, 2024. [Online]. Available: https://www.kaggle.com/datasets/cm037divya/pcos-dataset

[27] S. Sadhwani, U. K. Modi, R. Muthalagu, and P. M. Pawar, "SmartSentry: Cyber threat intelligence in industrial IoT," *IEEE Access*, vol. 12, pp. 34720–34740, 2024, doi: 10.1109/ACCESS.2024.3371996.

[28] M. Lokanan, "Exploring resampling techniques in credit card default prediction," *Research Square*, pp. 1–37, 2024, doi: 10.21203/rs.3.rs-4087259/v1.

[29] P. Sedgwick, "Pearson's correlation coefficient," *BMJ*, vol. 345, pp. e4483–e4483, 2012, doi: 10.1136/bmj.e4483.

[30] J. Sainui and C. Srivisal, "Unsupervised feature selection with least-squares quadratic mutual information," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 22, no. 3, pp. 1619–1628, 2021, doi: 10.11591/ijeecs.v22.i3.pp1619-1628.

[31] Y. Chen and M. L. Wong, "An ant colony optimization approach for stacking ensemble," in *2010 Second World Congress on Nature and Biologically Inspired Computing (NaBIC)*, 2010, pp. 146–151, doi: 10.1109/NABIC.2010.5716282.

[32] O. Peretz, M. Koren, and O. Koren, "Naive Bayes classifier – an ensemble procedure for recall and precision enrichment," *Engineering Applications of Artificial Intelligence*, vol. 136, 2024, doi: 10.1016/j.engappai.2024.108972.

[33] S. Ahmed *et al.*, "Exploring deep learning and machine learning approaches for brain hemorrhage detection," *IEEE Access*, vol. 12, pp. 45060–45093, 2024, doi: 10.1109/ACCESS.2024.3376438.
[34] H. Yun, "Prediction model of algal blooms using logistic regression and confusion matrix," *International Journal of Electrical and Computer Engineering*, vol. 11, no. 3, pp. 2407–2413, 2021, doi: 10.11591/ijece.v11i3.pp2407-2413.
[35] Y. Song, X. Meng, and J. Jiang, "Multi-layer perception model with elastic grey wolf optimization to predict student achievement," *PLoS ONE*, vol. 17, no. 12, 2022, doi: 10.1371/journal.pone.0276943.
[36] D. Berrar, "Cross-validation," in *Encyclopedia of Bioinformatics and Computational Biology*, 2019, pp. 542–545, doi: 10.1016/B978-0-12-809633-8.20349-X.

BIOGRAPHIES OF AUTHORS

Kazi Abu Taher received the Ph.D. degree from the Bangladesh University of Engineering and Technology, Dhaka, Bangladesh. He is currently a research fellow with the city, University of London, U.K., and a visiting scholar with the National University of Singapore. He is a Professor with the Bangladesh University of Professionals. His research interests include photonics, networking, communication security, block chain security, machine learning and ICT in pedagogy. He has been an expert reviewer of IEEE conferences, Elsevier, and Springer, journals. He has also been chair, technical program committee of several IEEE international conferences. He can be contacted at email: kataher@bup.edu.bd.

Samia Ahmed received the B.Sc. degree in information and communication engineering from the Bangladesh University of Professionals, in 2022, where she is currently pursuing the M.Sc. degree in ICE. She also served as a teaching assistant at Bangladesh University of Professionals. Her research interest includes healthcare with ML. In 2023, she was a recipient of the Chancellor's Gold Medal. She can be contacted at email: samiaomi22@gmail.com.

Jannatul Ferdous Esha received the B.Sc. degree in information and communication engineering from Bangladesh University of Professionals, in 2022, where she is currently pursuing the M.Sc. degree in ICE. Her research interests include the use of machine learning and vision for healthcare and medicine. She can be contacted at email: jannatul.ferdous.esha11235@gmail.com.

Md. Sazzadur Rahman received B.Sc. and M.S. degrees in applied physics, electronics, and communication engineering from University of Dhaka, Bangladesh, in 2005 and 2006, respectively, and Ph.D. degree in material science from Kyushu University, Japan, in 2015. From May 2009 to November 2018, he was faculty member with Faculty of Computer Science and Engineering, Hajee Mohammad Danesh Science and Technology University. Since November 2018, he has been with Institute of Information Technology, Jahangirnagar University, Dhaka. Currently, he is a Professor. His research interests include material science for computer applications, surface science, ubiquitous computing, WSNs, machine learning, and IoT. He can be contacted at email: sazzad@juniv.edu.

A. S. M. Sanwar Hosen earned his M.S. and Ph.D. degrees in computer science and engineering from Jeonbuk National University, South Korea. He is currently serving as Assistant Professor in Department of Artificial Intelligence and Big Data at Woosong University, South Korea. He has authored numerous articles in reputable journals and presented at international conferences. His research interests encompass wireless sensor networks, IoT, fog-cloud computing, cyber-security, artificial intelligence, and blockchain technology. Additionally, he has served as reviewer, guest editor for prestigious journals and as program committee member for several international conferences, including those organized by IEEE and ACM. He can be contacted at email: sanwar@wsu.ac.kr.