IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, April 2024

Explainable Artificial Intelligence based Detection and Early Diagnosis of Polycystic Ovaries Syndrome using Optimized Hybrid Deep Learning Technique

Amol Bajirao Kale, Preeti Baban Lokhande, Ramshi Purushottam Pathak, Shivaji Arun Shinde
MAEER'S MIT Arts, Commerce, and Science College, Alandi, Pune, Maharashtra, India

Abstract: *Customer satisfaction is directly related with the customer retention. The marketer should understand the needs and expectations of his customers for making an effective marketing strategy. Measurement of customer satisfaction enables the firm to deliver maximum value to the customer. Delivering the values to customers facilitates in the creation of loyal customers. The main thrust area among these challenges is the dissatisfaction of customers. The main reason behind this dissatisfaction is the expectations of modern customers who are tech-savvy guys. The digitalization in the area of business is likely to continue in future which will create more challenges before the marketers. Hence customer satisfaction cannot be ignored in the modern digital age*

Keywords: Customer Satisfaction, Digital Age, Entrepreneurship, Business

I. INTRODUCTION

1.1 Background of the Study

Polycystic ovary syndrome (PCOS) is a condition which affects the female's reproductive system and is caused by both genetic and biological factors. PCOS increases the complications in the female's reproductive system. Subsequently, medical professionals and researchers have described PCOS can have several impacts on a female's health by causing hormonal imbalances, presence of polycystic ovaries, and ovulatory dysfunction (Aggarwal et al., 2023) [1] (Tiwari et al., 2022) [2]. This can have a significant impact on the reproduction capacity and overall health. It can be inferred from the available information that 1 in 10 females suffer from PCOS (Farquhar et al., 2022) [3]. Considering the severity, it is important to treat PCOS in the early stage. In most common cases, this condition is identified through manual diagnosis using a transvaginal ultrasound. The adoption of ultrasound technology in characterizing the morphology of ovaries helps in the accurate diagnosis of the PCOS condition (Al Wattar et al., 2021) [4] (Teede et al., 2022) [5]. With the emergence of Machine Learning (ML) techniques have transformed the diagnosis process. Machine learning belongs to the class of Artificial Intelligence (AI) which automates the diagnosis process and provides high accurate results. However, the implementation of ML algorithms always comes with the complexity of the model and it is difficult to achieve a balanced tradeoff between the efficiency and complexity of the model. To overcome the drawbacks associated with complex models, this research emphasizes the implementation of Explainable Artificial Intelligence (XAI) for effective detection of the PCOS condition.

1.2 Rationale and Significance of Study

PCOS is one of the most common hormonal disorders, affecting one out of every ten premenopausal reproductive women globally. PCOS is connected with an abnormally high level of male androgen hormone in the female body, causing a chronic disturbance in hormonal levels and, as a result, negatively affecting normal ovarian processes, leading in the creation of many cysts inside the ovary. It is a life-long syndrome because the underlying cause is unknown. It can damage either one or both ovaries. Hyperandrogenism is a clinical indication of PCOS that causes facial hair development, ovarian cysts, and irregular menstrual cycles. Over the previous few decades, technology has transformed

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, April 2024

our environment and changed our lives, making them easier to live on a daily basis. There are numerous ways that modern technology is transforming humans. Machine learning, a research area which is under Artificial Intelligence that enables computers to learn without being explicitly instructed, is currently heavily used in the healthcare industry. Machine learning can handle enormous datasets, translate studied data into relevant therapeutic insights, and aid in the diagnosis of a wide range of disorders.

The concept of XAI is similar to that of traditional AI models but with the ability to interpret the results. XAI is a set of tools and frameworks which helps to understand and interpret the predictions made by conventional ML models. XAI models generate results which can be easily interpreted or understood by humans. Unlike traditional ML models, XAI consists of two different elements namely an explanatory module and explanation interface module. Although ML algorithms achieve high accuracy in all tasks, they lack explainability. The implementation of XAI provides high accuracy, effectiveness along with better explainability, transparency, and trustworthiness. As a result, there is a great significance of XAI in applications such as disease detection, prediction and classification. Considering the advantages of XAI, this research intends to detect PCOS using XAI tools. The study intends to explore the potentiality of XAI and investigates the performance efficiency of these tools in the disease detection process.

1.3 Statement of Problem

In recent times, ML and DL (Deep Learning) techniques are used to solve complex business problems and are applied in every field to develop intelligent solutions. Several DL models were employed in real time disease detection and diagnosis. All the while deep learning methods often achieve a high level of accuracy in picture classification, they have the constraint of requiring a lot of processing complexity and time to perform, which makes them difficult to use in practical applications. As a result, an integrated or extended ML-based strategy to predicting PCOS using image data may improve prediction performance while decreasing computing complexity.

Despite the availability of several decision support models to achieve improved accuracy in PCOS detection and diagnosis there is a lack of an effective model which can work effectively with a large scale and imbalanced dataset (Nsugbe, 2023) [6]. Existing works reveal that there is still a huge scope for improvement in the precision metric. The main limitations of ML and DL models are based around the sample size and its nature. Further, the effectiveness of these models in the clinical environment needs to be validated using a larger and more diverse sample set. Hence, it is important to explore more models for identifying PCOS using larger datasets using the concepts of XAI.

II. RELATED WORKS

A significant amount of research has been dedicated to identifying PCOS using AI models. Several research works have used different ML and DL algorithms for the accurate detection of PCOS (Lv et al., 2022) [7] (Makhdoomi et al., 2022) [8] (Bharati et al., 2022) [9]. However, the lack of explainability in ML and DL models motivates the researchers to look for more advanced techniques such as XAI. The work presented in (Elmannai et al., 2023) [10] employed XAI along with feature selection techniques using different ML models for detecting PCOS. ML algorithms such as logistic regression (LR), random forest (RF), decision tree (DT), naive Bayes (NB), support vector machine (SVM), k-nearest neighbor (KNN), XGboost, and Adaboost algorithm were evaluated for selecting optimal features from the clinical data. Stacked models which integrate ML models were employed for improving the performance. The ML models were optimized using Bayesian optimization and the problem of data imbalance was addressed using a SMOT technique. Results show that the stacking ML with RF based feature selection achieved a phenomenal accuracy of 100% compared to other techniques. The authors in (Khanna et al., 2023) [11] employed a distinctive XAI model for detecting PCOS. The multi-stack ML models achieved a high detection accuracy of 98%. Results show that XAI models exhibit better predictions with interpretable results and trustworthiness. Different XAI tools such as SHAP (SHapley Additive Values), LIME (Local Interpretable Model Explainer), ELI5, Qlattice along with RF classifier were studied. The risk factors of PCOS using the XAI model is studied by (ÇİÇEK et al., 2021) [12]. A LIME method was used to analyze the POCS condition and the data collected from the “Polycystic Ovary Syndrome” dataset was used to train the RF model for estimating the risk factors of PCOS. Results show that the RF classifier achieved a phenomenal accuracy, sensitivity, and specificity of 86.03%, 86.32%, and 85.37% respectively. An extended ML model for detecting PCOS using an ultrasound image of an ovary is presented in (Suha& Islam, 2022) [13]. A Convolutional Neural Network

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, April 2024

(CNN) is implemented for detecting PCOS which is trained and tested using different ultrasound images of ovaries. A transfer learning approach is used for extracting relevant features from the images and the extracted features were combined using an ensemble ML technique. Fundamental base models along with bagging and boosting ensemble techniques were used for reducing the dimensionality of the feature set. This enabled the model to distinguish the images of PCOS and non-PCOS ovaries. The proposed approach improved the accuracy and also reduced the execution time and training time required for training the model, compared to other state of art techniques. The best results were achieved by employing the VGGNet 16 pre-trained model along with CNN model. The VGGNet 16 model was used as a feature extractor and the XGBoost model was used as a classifier which achieved a classification accuracy of 99.89%. It can be observed from the literary works that the deployment of XAI tools along with ML and DL models can enhance the performance of PCOS detection.

III. AIMS AND OBJECTIVES

The preliminary aim of this research is to automatically detect PCOS using the concept of XAI. The objectives of this research are as follows:

3.1 Research Objectives

- To implement a hybrid deep learning model for detecting and diagnosing PCOS.
- To incorporate XAI tools for achieving better interpretability, explainability, transparency and trustworthiness.
- To increase the quality of the input images, and enhance the image attributes in order to make them more appropriate for the classification process.
- To achieve a high classification accuracy and better precision with lesser error rate.
- To validate the performance of the proposed approach by comparing the results with other similar existing models.

IV. METHODOLOGY AND TECHNIQUES USED

This research aims to design an ensemble learning framework for detecting PCOS from the ultrasound ovary images. The proposed ensemble learning model will be developed using a deep neural network (DNN) combined with a random forest (RF) algorithm. The DNN-RF model will leverage the advantages of both DNN and RF algorithms for performing complex tasks such as disease detection and classification.

The proposed framework will consist of two modules; (a) disease detection and classification using the DNN-RF model and (b) explanation of the output generated. In the first part, a RF model will be used for selecting features and the DNN model will be used for classifying PCOS as PCOS and non-PCOS ovaries. The model will be trained using a larger dataset in order to achieve a better performance. The DNN-RF model will be optimized using a Bayesian optimization model. In the second part, two different XAI tools will be implemented for explaining or interpreting the model's performance. Accuracy and traceability are two important attributes of XAI models and in this content, this research proposes the implementation of a LIME-Local Interpretable Model-Agnostic Explanations, and DeepLIFT (Deep Learning Important Features) tools (Lundberg & Lee, 2017) [14]. The LIME method explains the performance of the ML classifiers in terms of classification and disease detection and DeepLIFT compares the activation of each neuron to its reference neuron and shows a traceable link between each activated neuron and even shows dependencies between them. This work mainly focuses on the explanation part and aims to address the challenges associated with the implementation of XAI and contribute to the existing research gap.

The techniques used in this research are as follows:

- Dataset: PCOS dataset (Elmannai et al., 2023) [10], and Kaggle dataset
- Techniques for PCOS detection and classification:Deep Neural Network (DNN) and Random Forest (RF) algorithm.
- Optimization Technique: Bayesian Optimization
- XAI Tools: LIME and DeepLIFT
- Performance Evaluation Metrics: Accuracy, Precision, Recall, and F1 score
- Software Tools: MATLAB, PYTHON, and Graphic Processing Units (GPU)

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, April 2024

V. EXPECTED OUTCOME

The proposed research aims to design an effective model for detecting PCOS using a hybrid DNN-RF approach, leveraging the advantage of XAI tools. It is expected that the implementation of XAI will significantly help in achieving desired PCOS detection performance along with better explainability and interpretability. The proposed DNN-RF model with XAI tools is expected to provide better results compared to other existing models. It can also be expected that the proposed model will overcome the complexities associated with conventional ML models in terms of improving the computational time and efficiency and will be trained to deliver effective results with high classification and detection accuracy.

REFERENCES

[1]. Aggarwal, N., Shukla, U., Saxena, G. J., Kumar, M., Bafila, A. S., Singh, S., &Pundir, A. (2023). An Improved Technique for Risk Prediction of Polycystic Ovary Syndrome (PCOS) Using Feature Selection and Machine Learning. In *Computational Intelligence: Select Proceedings of InCITe 2022* (pp. 597-606). Singapore: Springer Nature Singapore.

[2]. Tiwari, S., Kane, L., Koundal, D., Jain, A., Alhudhaif, A., Polat, K., ... &Althubiti, S. A. (2022). SPOSDS: A smart Polycystic Ovary Syndrome diagnostic system using machine learning. *Expert Systems with Applications*, *203*, 117592.

[3]. C. Farquhar, Introduction and history of polycystic ovary syndrome, in: GT Kovacs, R Norman (Eds.), Polycystic Ovary Syndrome [Internet], second ed., Cambridge University Press, 2001, pp. 4–24, Available from: https://www.cambridge.org/core/product/identifier/CBO9780511545191A008/ type/book_part [cited 2022 Oct 31].

[4]. Al Wattar, B. H., Fisher, M., Bevington, L., Talaulikar, V., Davies, M., Conway, G., &Yasmin, E. (2021). Clinical practice guidelines on the diagnosis and management of polycystic ovary syndrome: a systematic review and quality assessment study. *The Journal of Clinical Endocrinology & Metabolism*, *106*(8), 2436-2446.

[5]. Teede, H. J., Garad, R. M., Melder, A., Norman, R. J., & Boyle, J. (2022). Letter to the Editor from Teede:"Clinical Practice Guidelines on the Diagnosis and Management of Polycystic Ovary Syndrome: A Systematic Review and Quality Assessment Study". *The Journal of Clinical Endocrinology & Metabolism*, *107*(3), e1321-e1322.

[6]. Nsugbe, E. (2023). An artificial intelligence-based decision support system for early diagnosis of polycystic ovaries syndrome. *Healthcare Analytics*, *3*, 100164.

[7]. Lv, W., Song, Y., Fu, R., Lin, X., Su, Y., Jin, X., ... & Huang, G. (2022). Deep learning algorithm for automated detection of polycystic ovary syndrome using scleral images. *Frontiers in Endocrinology*, *12*, 1869.

[8]. Makhdoomi, A., Jan, N., Palak, P., &Goel, N. (2022, December). Machine learning techniques for medical images in PCOS. In *2022 4th International Conference on Artificial Intelligence and Speech Technology (AIST)* (pp. 1-6). IEEE.

[9]. Bharati, S., Podder, P., Mondal, M. R. H., Surya Prasath, V. B., & Gandhi, N. (2022, March). Ensemble Learning for Data-Driven Diagnosis of Polycystic Ovary Syndrome. In *Intelligent Systems Design and Applications: 21st International Conference on Intelligent Systems Design and Applications (ISDA 2021) Held During December 13–15, 2021* (pp. 1250-1259). Cham: Springer International Publishing.

[10]. Elmannai, H., El-Rashidy, N., Mashal, I., Alohali, M. A., Farag, S., El-Sappagh, S., &Saleh, H. (2023). Polycystic Ovary Syndrome Detection Machine Learning Model Based on Optimized Feature Selection and Explainable Artificial Intelligence. *Diagnostics*, *13*(8), 1506.

[11]. Khanna, V. V., Chadaga, K., Sampathila, N., Prabhu, S., Bhandage, V., &Hegde, G. K. (2023). A Distinctive Explainable Machine Learning Framework for Detection of Polycystic Ovary Syndrome. *Applied System Innovation*, *6*(2), 32.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, April 2024

[12]. ÇİÇEK, İ. B., KÜÇÜKAKÇALI, Z., & YAĞIN, F. H. (2021). Detection of risk factors of PCOS patients with Local Interpretable Model-agnostic Explanations (LIME) Method that an explainable artificial intelligence model. *The Journal of Cognitive Systems*, *6*(2), 59-63.

[13]. Suha, S. A., & Islam, M. N. (2022). An extended machine learning technique for polycystic ovary syndrome detection using ovary ultrasound image. *Scientific Reports*, *12*(1), 17123.

[14]. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, *30*.