International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

PCOScare: Conventional Machine Learning Classifiers for Diagnosing and Prevention

Vaibhav C. Gandhi[1], Dr. Khyati R. Nirmal[2], Dr. Uma Maheswari[3], Dr. Sudha Rajesh[4], Dr. P. Tharcis[5], Dhruvi Thakkar[6]

Submitted: 29/01/2024 **Revised**: 07/03/2024 **Accepted**: 15/03/2024

Abstract: PCOS is a common endocrine disorder that affects women of reproductive age. Timely identification and diagnosis of PCOS are essential for the successful treatment and prevention of related health complications. ML techniques have shown promise in automating PCOS diagnosis using various clinical and biochemical features. Though, the presentation of these mockups seriously relies on the selection of relevant features, as including irrelevant or redundant features can lead to overfitting and decreased generalization performance. In this study, we propose an optimized feature selection approach for PCOS detection using ML algorithms. We first compile a comprehensive dataset comprising clinical and biochemical features commonly associated with PCOS, including hormone levels, menstrual irregularities, and anthropometric measurements. For the purpose of handling missing data and scaling features properly, feature preprocessing techniques like normalization and imputation are used. We next investigate several feature selection strategies, such as filter, wrapper, and embedding approaches, to find the most relevant characteristics for PCOS identification. We use cross-validation and presentation needles with accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC) in a methodical review approach to maximize feature selection. The presentation of respectively feature assortment method is compared to assess its effectiveness in identifying discriminative features for PCOS diagnosis. Furthermore, we investigate the impact of feature selection on different ML algorithms, including support vector classifier, random forests, and Xg-boosting classifiers. Our results demonstrate that feature selection significantly improves the performance of PCOS detection models by reducing dimensionality and focusing on the most relevant features. Moreover, we identify key features that contribute most to the discriminative power of the models, providing insights into the underlying characteristics of PCOS. The optimized feature selection approach proposed in this study offers a promising strategy for developing accurate and interpretable ML models for PCOS detection, ultimately aiding clinicians in early diagnosis and personalized treatment planning.

Keywords: *Polycystic Ovary Syndrome, Random Forest Classifier, Machine Learning, Early Detection, Prevention*

1. Introduction

With an estimated frequency of 6% to 15% worldwide, polycystic ovarian syndrome, or PCOS, is one of the most prevalent endocrine illnesses affecting women of reproductive age. Menstrual abnormalities, hyperandrogenism, and polycystic ovarian morphology on ultrasonography are among the diverse symptoms that define it and infertility.

To lessen the negative effects of PCOS on health and enhance the lives of those who are affected, early detection and efficient treatment are crucial. However, PCOS diagnosis can be challenging and frequently necessitates a thorough assessment of imaging results, biochemical markers, and clinical symptoms. Traditional diagnostic criteria, such as those outlined by the Rotterdam criteria and the Androgen Excess Society criteria, rely on the presence of specific clinical and biochemical features. [1,2,3,4] While these criteria have been valuable for guiding clinical diagnosis, they may lack sensitivity and specificity, leading to under diagnosis or misdiagnosis of PCOS.

Hormone levels are outside of balance in women with PCOS. It causes main wellbeing matters, such as irregular periods and difficulty becoming pregnant. A somewhat smaller percentage of African Americans—8%—and White Americans—4.8%—are impacted via PCOS. In Spain, 6.8% is found, but in Asia, 31.3% is found. PCOS prevents the follicle from developing normally, which is the main root cause of infertility. Furthermore, 30–70% of females with PCOS are overweight. The symptoms of PCOS include irregular menstruation, high levels of testosterone creation, and a noteworthy ovarian cyst prevalence. Figure 1 demonstrates the disparity amongst a polycystic ovary and a normal ovary [11].

[1]Department of Computer Engineering, Madhuben and Bhanubhai Patel Institute of Technology (MBIT), CVM University, Anand, Gujarat – India.
[2]Department of Computer Engineering, SNJB's Late Sau. KBJ College of Engineering, Neminagar, Chandwad – 422011, India.
[3]Division of Data Science and Cyber Security, Karunya Institute of Technology and Sciences, Karunya Nagar, Coimbatore, Tamil Nadu – 641114, India.
[4]Department of Computational Intelligence, College of Engineering and Technology, School of Computing, SRMIST, Kattankulathur, Chennai-India.
[5]Department of Electronics and Communication Engineering, SRM Madurai College for Engineering and Technology, Pottapalayam, India.
[6]Department of Computer Science & Engineering, Drs Kiran & Pallavi Patel Global University (KPGU University), Vadodara, Gujarat – India.

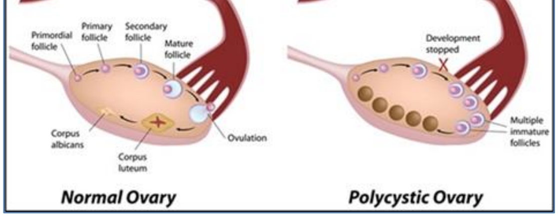

Fig. 1. The Different Patterns in Ovaries

Technology has become an integral part of human existence, influencing healthcare trends and providing easy access to diagnostic tools. Advancements in technology have enabled ML to learn from available information, aiding in illness prediction, particularly in clinical decision support. This technology has become an essential part of our daily lives. Healthcare services may now be automated to diagnose illnesses based just on their symptoms thanks to the advancements in AI and ML. The study of PCOS diagnosis has lately piqued the interest of several academics. Physicians will be able to evaluate patients more quickly with an early diagnosis. Evaluating the most modern methods for PCOS detection is the primary objective of this research study [13]. PCOS women are categorized using hybrid ML methods that include binary classification.

The emergence of ML techniques offers new opportunities for enhancing PCOS diagnosis through automated analysis of patient data. By leveraging large datasets containing a multitude of clinical and biochemical features, ML models can potentially identify patterns and associations that are not readily apparent to human observers. Moreover, these models have the potential to improve diagnostic accuracy, enable personalized risk assessment, and facilitate timely intervention for individuals at risk of developing PCOS-related complications [13, 16, 17]. However, the successful development and deployment of ML models for PCOS detection hinge on several critical factors, including feature selection, model interpretability, and generalizability. Feature selection plays a pivotal role in identifying the most informative and discriminative features from the vast array of available variables, thereby reducing dimensionality, mitigating overfitting, and enhancing model performance. Moreover, selecting relevant features can improve the interpretability of the models by identifying the underlying biological mechanisms and clinical factors associated with PCOS.

In this research paper, we present an optimized feature selection approach for developing ML models for PCOS detection. We aim to investigate various feature selection methods and evaluate their effectiveness in identifying discriminative features associated with PCOS. Furthermore, we explore the impact of feature selection on different ML algorithms and assess the performance of the resulting models using rigorous evaluation metrics. The following are the article's primary objectives:

- The paper's principal objective is to create a PCOS initial identification model. The probability of continuing issues is decreased by early PCOS identification. ML is used in many instances to construct the suggested layering ensemble ML framework.
- ML classifiers like LR, SVM, RF, Decision, and Extra Trees are reviewed for effectiveness, and a combined arranging of meta-learners is utilized to generate a more precise and trustworthy classifier.
- To identify the data patterns that are most contributing to PCOS illness, PCOS exploratory data analysis, or PEDA, is carried out. Graphs, charts, and statistical evaluation of information form the foundation of the PEDA.
- Using the feature selection approach, identify the most significant characteristics of those with PCOS according to the provided dataset.

2. Literature Survey

This section reviews the literature that is relevant to the recommended research investigation. Analysis is done on previous applications of cutting-edge research for PCOS prediction. Analyzed are the suggested methods and associated outcomes of the study.

V.V.Khanna et al.[1] developed an AI approach using heterogeneous ML and deep learning classifiers to predict PCOS among prolific patients. They used an open-source dataset of 541 patients from Kerala, India. Logical AI strategies, such as SHAP, LIME, ELI5, Qlattice, and Irregular Timberland, were used to make the predictions justifiable, interpretable, and reliable. The goal is to accurately distinguish PCOS in patients while proposing a robotized screening design with logical ML devices to help restorative experts in decision-making. This study contributes to the advancement of healthcare and the development of new treatments for PCOS.

According to S. Nasim et al. [2], a common endocrine condition that affects women globally and causes infertility in 4,444 of them. Moreover, it results in insulin resistance, which raises the risk of diabetes. Higher levels of androgens in PCOS-afflicted women might result in irregular menstrual periods, acne, thinning hair, and excessive hair. In 2020, a comprehensive study involving 4,444 women in India found that 16,444% of them had PCOS. It can cause major health problems like infertility, cancer, heart disease, and acne scarring if left untreated. Because of the stigma associated with the diagnosis and the protracted first diagnostic phase, the majority of women with PCOS go undetected despite the condition's long-term and potentially fatal consequences.

The focus of S. Bharati et al. [3] is on the data-driven diagnosis of female PCOS. This entails using the Kaggle repository's public datasets in conjunction with her machine learning method. 43 characteristics from her 4,444,541 women, including 177 PCOS patients, are included in this dataset. To identify the best characteristics for PCOS prediction, we first use a univariate feature selection approach. After a ranking of the features is determined, the ratio of luteinizing hormone (LH) to follicle stimulating hormone (FSH) is revealed to be the most significant attribute. The dataset is then subjected to holdout and cross-validation techniques to distinguish training from testing data. The dataset is subjected to a number of classifier applications, such as logistic regression, hybrid random forest, random forest, gradient boosting, and logistic regression (RFLR). The greatest test precision of 91.01% and recall of 90% are also demonstrated by the results when 40-fold cross-validation is applied to the top 10 features for RFLR. As a result, PCOS patients may be accurately classified using RFLR.

S. Prasher et. al. [4] highlight hormone-related conditions like PCOS, which include irregular periods, hair growth, weight gain, skin discoloration, diabetes, and infertility. This study aims to predict PCOS prevalence using classification techniques and compare algorithms' performance. Moreover, it utilizes a relationship-based feature-selection technique along with Engineered Minority Oversampling Technique (Destroyed) for developing a classifier. Henceforth, examine seven different existing ML approaches namely: SVM, multilayer SVM linear bit with calculated regression, RF, LR, DT, KNN and ANN on classification accuracy basis for included chosen features. The performance estimation is carried out based on the Kaggle repository’s dataset for PCOS analysis using classifiers provided in Weka toolkit [16]. In terms of accuracy rate SVC and ensemble-based RF classifiers have 96.22% and 98.89% respectively during evaluation process

S. Alshakrani et.al. [5] This study aims to understand PCOS, a hormonal illness that moves a female's regarding ability and increases the risk of long-term complications. Primary discovery is crucial, and ensemble learning techniques and ML are used to enhance detection powers. The use of hybrid machine learning models, such as the Hybrid Random Forest Logistic Regression,

Extreme Boosting with Random Forest, Linear Support Vector Machine, Light Gradient Boosting Model, and Cat Boost model. The PCOS dataset, recovered since the Kaggle source, includes clinical and physical characteristics of women from ten clinics in Kerala, India. The top 14 clinical and physical criteria are chosen to identify PCOS, with follicle numbers being the most important. Cat Boost model outperforms extra models with an accurateness score of 92%.

Prajna K B [6], defined hormonal disorder that affects women who are fertile and can result in weight gain, irregular menstruation periods, acne, facial hair development, and hair loss. Treating PCOS in its early stages might be difficult. We suggest a method that use ML techniques to predict and identify PCOS using the fewest possible parameters. After consulting with gynecologists, we employed a dataset from the open-source database "KAGGLE" to determine the top 10 to 15 attributes. The model was trained, validated, and tested using four ML algorithms: the Chi-Square algorithm, RF classifier, logistic regression, and DT classifier. According to our findings, the Random Forest Classifier (Chi-Square) outperforms the other algorithms in terms of accuracy.

The prevalent cause of female infertility, which affects a sizable percentage of females of procreant age, was reported by M. Inan et al. Because of XGBoost's strong detection capabilities, the authors recommend utilizing it for early PCOS identification. They resampled data using SMOTE and ENN to overcome class imbalance and data outliers. They identified 23 metabolic and clinical characteristics that best characterize PCOS symptoms using statistical correlation techniques. In identifying patients without PCOS, the XGBoost classifier beat all other classifiers, with a 98% recall and a 10-Fold Cross-validation score of 96.03%. To verify the model's efficacy, it was run on a benchmark dataset from Kaggle [7].

S. Vedpathak et. al. [8] where a medical illness known as PCOS affects women who are fertile and results in hormonal disorders. The hormonal imbalance causes a missing or delayed menstrual cycle. The primary signs of PCOS in females include irregular periods, excessive weight gain, acne, hair loss, facial hair growth, and skin discoloration. In rare cases, PCOS can also lead to infertility. The existing methodologies and medications are inadequately for early-stage discovery and prediction. To handle this issue, we propose an approach that can offer assistance with early determination and PCOS treatment forecast given an ideal and negligible set of criteria. To find out whether a lady has PCOS, five distinctive machine learning classifiers have been utilized: Irregular Timberland, SVM, Calculated Relapse, Gaussian Naïve Bayes, and K Neighbors. Out of the 41 highlights added up, the 30 best highlights from the dataset were recognized and included to the highlight vector utilizing the Chi Square strategy. Furthermore, we assessed each classifier's results and found that the arbitrarily chosen Woodland Classifier had the highest and most dependable exactness. You'll utilize the Prasoon Kottarathil-owned dataset on Kaggle for preparing and testing.

H. Elmannai [9] explain about the global health issue affecting women worldwide. This investigate intentions to improve early detection and treatment of Type 2 and gestational diabetes using machine learning models like relapse, choice tree, naïve bayes, bolster vector machine, k-nearest neighbor, and Adaboost calculation. It is recommended to stack ML models using REF feature selection to improve performance. According to experimental data, the maximum accuracy of 100 was obtained by stacking ML with REF.

3. Proposed Methodology

In this comprehensive research endeavour, our methodology for predicting PCOS diagnosis has been meticulously crafted to ensure the accuracy and reliability of the predictive models utilized. Figure 2 below shows the phases of prediction PCOS.

Table 1. Summary of Literature Survey

Works	Methodology	Dataset	High Performance Params.	Limitations
1	Explainable AI (XAI), SHAP, LIME, ELI5,	Formally shared dataset	Accuracy, Precision, Recall, and F1-score	The scope of the study was constrained by the availability and quality of the data as well as geographical variations in healthcare practices. The model's effectiveness may also be impacted by changes to healthcare laws and advancements in diagnostic standards.
2	Feature selection and RF classifier	Formally shared dataset	Chi-Square Method, Extra Tree Classifier, Correlation Matrix	NA
3	Hybrid RFLR	Formally shared dataset	Accuracy Recall	The study employs RFLR and other algorithms with fewer characteristics and resulting in less computing time for their classification and other techniques.
4	Traditional ML techniques, SMOT and Feature selection	Formally shared dataset	Accuracy	Modifications to healthcare legislation and improvements in diagnostic criteria may also have an influence on the model's efficacy
5	Hybrid ML Model	Formally shared dataset	Accuracy, Precision, Recall, F1-	A little amount of tabular data from Kerala,

			Score, ROC curve plot, K – fold Cross Validation	India, was used in the present investigation.
6	RF classifier	Formally shared dataset	Accuracy Chi Square Method	The effectiveness of the approach may also be impacted by changes to healthcare laws and advancements in diagnostic standards.
7	SMOT, ENN Techniques for sampling, Extreme gradient Boosting classifier	Formally shared dataset	Accuracy Chi Square Method	NA
8	Traditional ML classifier	Formally shared dataset	Cross Validation Re call	For better performance, we would want to carry out more thorough hyperparameter tweaking of ML algorithms and better feature selection.
9	SMOT, ENN Techniques used for class imbalance and Stacking ML approach.	Formally shared dataset	Accuracy, Precision, Re Call, F score, Auc	NA

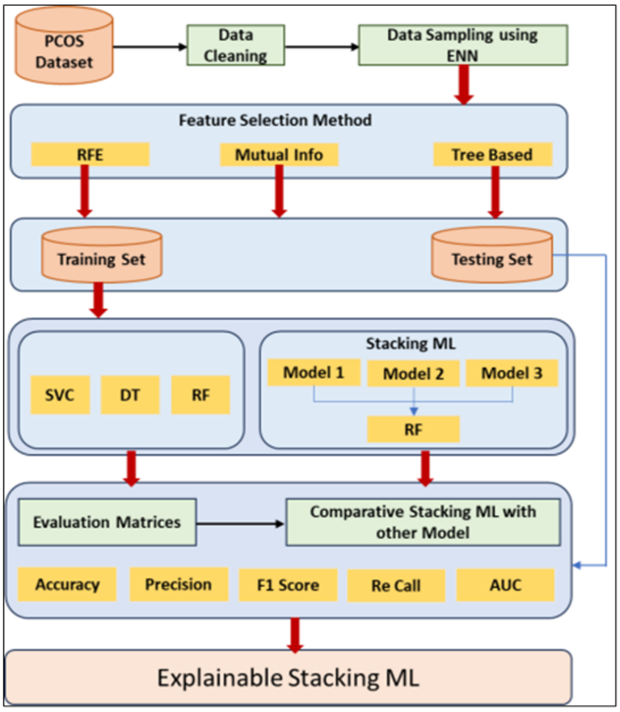

Fig. 2. Phases of Predicting Model

3.1. Dataset Polycystic ovary syndrome (PCOS):

Our dataset is derived from an extensive collection of clinical data obtained from 10 distinct hospitals across Kerala, India. This dataset encompasses a rich array of clinical features and symptoms associated with PCOS, capturing the diverse manifestations of this complex endocrine disorder. The dataset was split into two parts for ML purposes: 80% for training the model and 20% for testing its performance. This enables the model to learn from most of the data and then apply ML best practices to assess its performance on data that hasn't been seen before. The distribution of training and testing data across different categories is displayed in Table 3 below [3,4,5,6].

Table 2. Dataset Description

Various Level	*Training*	*Testing*
PCOS(Y)	434	109
PCOS(N)	434	109

3.2. Data Processing

A critical initial step in our methodology involves rigorous data preprocessing procedures aimed at enhancing data quality and suitability for predictive modelling. This encompasses various tasks, including the merging of relevant information from disparate sources, meticulous handling of missing values through imputation or removal strategies, and strategic feature engineering to transform and clean data types. Furthermore, to ensure uniformity and comparability across features, normalization techniques such as Min-Max scaling are applied, facilitating the standardization of feature scales and improving the convergence of our predictive models. Figure 1 below displays the heat map for the correlation features in the dataset.

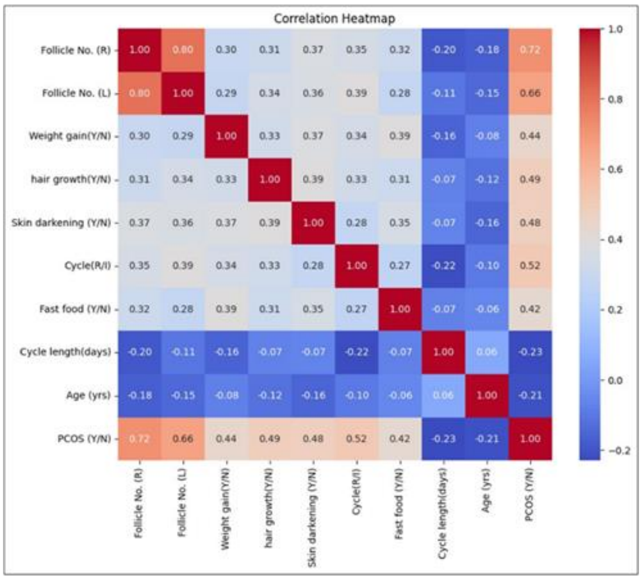

Fig. 3 The heat map highlights the dataset's correlation characteristics.

3.3. Feature Selection

Feature selection is a pivotal aspect of our methodology, enabling us to discern the most salient variables pertinent to PCOS diagnosis from the multitude of available features. Through a comprehensive analysis, we identify key indicators such as follicle count, weight gain, hair growth, skin darkening, cycle details, and age, which collectively contribute to the accurate prediction of PCOS. The figure 2 below distribution shows the distribution of the data frequencies in two categories: Yes & No

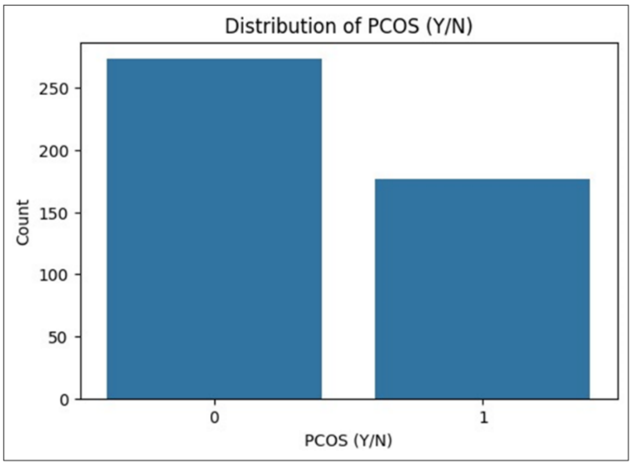

Fig. 4. Data Distribution

3.4. Handling Class Imbalance

Addressing class imbalance is paramount to the success of our predictive models. To mitigate the potential biases introduced by class imbalance, we employ the Edited

Nearest Neighbors (ENN) technique. By under sampling the majority class while preserving the overall distribution of the target variable, ENN ensures a balanced representation of PCOS cases, thereby enhancing the generalizability and effectiveness of important feature show in figure 5 in Random Forest Classifier.

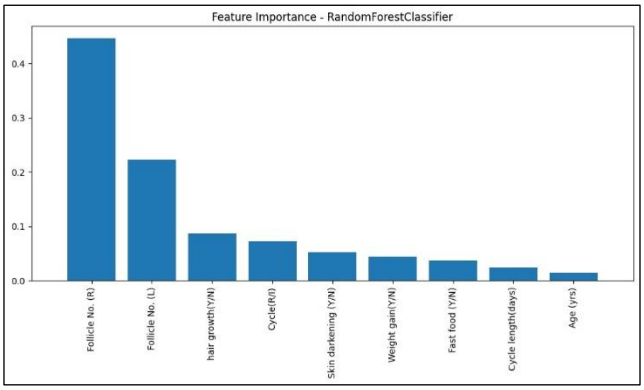

Fig. 5. Feature Importance: Random Forest Classifier

3.5. Modelling Approach

Our proposed methodology encompasses the application of two distinct yet complementary classification algorithms: the RF Classifier and the XGBoost Classifier. These algorithms are independently deployed to predict PCOS based on the identified features. Additionally, recognizing the potential synergies offered by ensemble learning, we introduce a Stacking Classifier. This innovative approach amalgamates the predictions of the base estimators—Random Forest and XGBoost—leveraging their unique strengths to enhance the overall predictive performance. Figure 6 below provides a detailed visualization of a decision tree within the Random Forest ensemble, featuring Gini impurity values at various nodes. The Gini impurity serves as a measure of the model's decision-making process, highlighting key points where feature values are evaluated for optimal classification which is shown in the below figure 6 decision tree visualization.

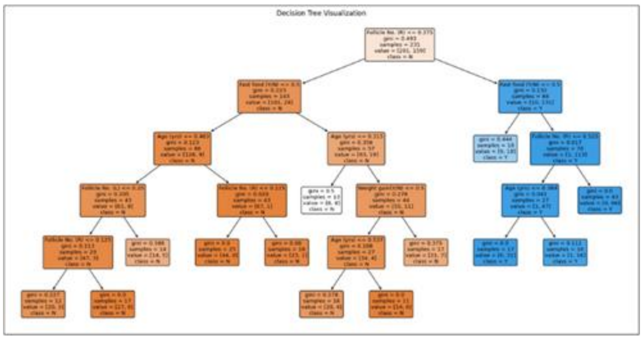

Fig 6. Decision Tree Visualization

4. Experimental Analysis

In this section, the following analysis shows the experimental results, and provide insights of graphs into the implications of our analysis for the PCOS. Experiment configuration is as follows:

- ❖ Hardware: AMD Ryzen 7, 1.90 GHz (16 GB RAM)
- ❖ Software: Google Colab CPU, T4 GPU
- ❖ Libraries: Matplotlib, Seaborn, Scikit-learn
- ❖ Architectures used: Random Forest Classification
- ❖ Dataset: PCOS

Three basic classification measures were used to appraise the representation's demonstration: accuracy, precision, and recall. Recall rated the model's capacity to identify positive instances, accuracy scored overall correctness, and precision assessed the significance of positive predictions. To upgrade show appraisal and decrease overfitting dangers, a strict 5-fold cross-validation method was utilized amid the hyperparameter alteration arrange. To do this, the dataset was partitioned into five subsets. The show was at that point prepared on four of the subsets, and its execution was surveyed on the fifth subset. To offer a more dependable assess of the model's execution, this strategy was carried out five times, with the normal results being calculated

The dataset underwent thorough preprocessing steps, including handling missing values, fixing data types, and selecting relevant features. Class imbalance was addressed using the Edited Nearest Neighbors (ENN) sampling strategy to ensure a balanced representation of PCOS and non-PCOS instances. Features were scaled using Min-Max scaling to standardize their range, facilitating smoother convergence during model training. The dataset was at that point part into preparing and testing sets, with 80% of the information utilized for preparing and 20% for testing to assess demonstrate execution on concealed information. Three different classifiers were trained and evaluated: Random Forest Classifier, XGBoost Classifier, and Stacking Classifier. The Random Forest Classifier utilized 150 estimators, employing the square root of features and a minimum of 10 samples per leaf to mitigate overfitting. Similarly, the XGBoost Classifier was configured with specific hyperparameters to optimize performance. Table 3 Classifier's accuracy below shows the accuracy of various classifier algorithm.

Table 3. Classifier 'S Accuracy

Classifier Algorithm	Accuracy
Random Forest	95%
XGBoost	56%
Stacking	94%
Cross Validation	92%

Additionally, a Stacking Classifier was constructed by combining predictions from the Random Forest and XGBoost models to improve predictive accuracy. The performance of the Stacking Classifier was assessed using cross-validation with 5 folds, providing an average accuracy score across different subsets of the data. A brief overview of the True, False, False and False forecasts gives important information about the model's performance and error rates.

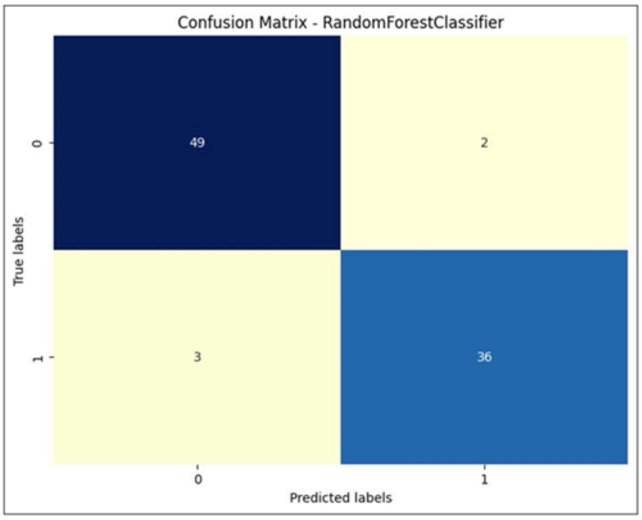

Fig 7. Normalized Confusion matrix of Random Forest Classifier

Finally, model performance was evaluated using various metrics such as accuracy scores, classification reports, and cross-validation scores. These metrics provided insights into the models' capabilities in predicting PCOS and ensured a thorough evaluation of their efficacy in clinical applications. The precision, recall and f1 score for each category is displayed in the figure 7 about random forest classifier.

Classification Report for Random Forest Classifier:

	precision	recall	f1-score	support
0	0.94	0.96	0.95	51
1	0.95	0.92	0.94	39
accuracy			0.94	90
macro avg	0.94	0.94	0.94	90
weighted avg	0.94	0.94	0.94	90

Fig 8. Random Forest Classifier Report

5. Discussion & Conclusion

We have to think ahead of problems associated with PCOS globally and regionally impact on women's health and quality of life, including its association with infertility, irregular menstrual cycles, obesity, insulin resistance, and psychological distress. Role of hormonal imbalances, such as hyperandrogenism and insulin resistance, in the development of PCOS. Recent research findings on genetic predispositions and environmental factors. Challenges associated with diagnosing PCOS, particularly due to its heterogeneous presentation and overlap with other conditions. Diagnostic criteria proposed by different organizations, such as the Rotterdam criteria and the Androgen Excess and PCOS Society criteria.

As Random Forest with hyperparameter tuning and with proper feature selection is giving 95% accuracy then other classifiers, examine the long-term health risks associated with PCOS, such as increased risk of type 2 diabetes, cardiovascular disease, endometrial cancer, and mental health disorders. Highlight the importance of early intervention and comprehensive management strategies. psychosocial implications of PCOS, including its effects on body image, self-esteem, and mental health. Address the challenges faced by women with PCOS in terms of stigma, social support, and access to healthcare. importance of continued research efforts in this area to improve the lives of individuals affected by PCOS and to reduce the burden of this condition on public health.

6. References and Footnotes

6.1. References

Author contributions

Vaibhav C. Gandhi: Conceptualization, Proposed Methodology, Implementation study **Dr. Khyati R. Nirmal**: Data curation, Writing-Original draft preparation, **Dr. Uma Maheswari**: Literature Survey, Validation., **Dr. Sudha Rajesh** Visualization, Investigation, Data Analysis **Dr. P Tharcis**: writing—review, editing and supervision. **Dhruvi Thakkar:** Implementation and Result. All authors contributed to the article and approved the submitted version.

Conflicts of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be

construed as a potential conflict of interest.

References

[1] "A Comparative Study for Detecting Polycystic Ovary Syndrome Using A Machine Learning Framework," IEEE Conference Publication | IEEE Xplore, Dec. 13, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10441348

[2] "Accessible Polycystic Ovarian Syndrome Diagnosis Using Machine Learning," IEEE Conference Publication | IEEE Xplore, May 27, 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9824049

[3] "Implementation of Various Machine Learning Algorithms to Predict Polycystic Ovary Syndrome," IEEE Conference Publication | IEEE Xplore, May 26, 2023. [Online]. Available:

https://ieeexplore.ieee.org/document/10170497

[4] “A Comparative Study on Different Machine Learning Algorithms to Detect PCOS,” IEEE Conference Publication | IEEE Xplore, Jul. 06, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10307174

[5] Majumdar, Sounak & Biswas, Saroj & Purkayastha, Biswajit. (2023). Diagnosis of Polycystic Ovary Syndrome Using Machine Learning Models: A Comparative Analysis. 1-5. 10.1109/i-PACT58649.2023.10434903.

[6] “A Comparative Study for Detecting Polycystic Ovary Syndrome Using A Machine Learning Framework,” IEEE Conference Publication | IEEE Xplore, Dec. 13, 2023. https://ieeexplore.ieee.org/document/10441348

[7] “Hybrid Machine Learning Algorithms for Polycystic Ovary Syndrome Detection,” IEEE Conference Publication | IEEE Xplore, Oct. 25, 2022. https://ieeexplore.ieee.org/document/10041525

[8] Elmannai H, El-Rashidy N, Mashal I, Alohali MA, Farag S, El-Sappagh S, Saleh H. Polycystic Ovary Syndrome Detection Machine Learning Model Based on Optimized Feature Selection and Explainable Artificial Intelligence. Diagnostics (Basel). 2023 Apr 21;13(8):1506. doi: 10.3390/diagnostics13081506. PMID: 37189606; PMCID: PMC10137609.

[9] Laganà AS, Vitale SG, Noventa M, Vitagliano A. Current Management of Polycystic Ovary Syndrome: From Bench to Bedside. Int J Endocrinol. 2018 Nov 14;2018:7234543. doi: 10.1155/2018/7234543. PMID: 30538744; PMCID: PMC6261402.

[10] D. Hdaib, N. Almajali, H. Alquran, W. A. Mustafa, W. Al-Azzawi and A. Alkhayyat, "Detection of Polycystic Ovary Syndrome (PCOS) Using Machine Learning Algorithms," 2022 5th International Conference on Engineering Technology and its Applications (IICETA), Al-Najaf, Iraq, 2022, pp. 532-536, doi: 10.1109/IICETA54559.2022.9888677.

[11] Vedpathak, Shreyas & Thakre, Vaidehi. (2020). PCOcare: PCOS Detection and Prediction using Machine Learning Algorithms. Bioscience Biotechnology Research Communications. 13. 240-244. 10.21786/bbrc/13.14/56.

[12] Prasher S, Nelson L, Sharma A. Evaluation of Machine Learning Techniques to Diagnose Polycystic Ovary Syndrome Using Kaggle Dataset. InInternational Conference on Emerging Trends in Expert Applications & Security 2023 Feb 17 (pp. 279-287). Singapore: Springer Nature Singapore.

[13] Subha R, Nayana BR, Radhakrishnan R, Sumalatha P. Computerized diagnosis of polycystic ovary syndrome using machine learning and swarm intelligence techniques.

[14] S. Ahmed et al., "A Review on the Detection Techniques of Polycystic Ovary Syndrome Using Machine Learning," in IEEE Access, vol. 11, pp. 86522-86543, 2023, doi: 10.1109/ACCESS.2023.3304536.

[15] P. Chitra, M. Sumathi, K. Srilatha, F. V. Jayasudha and S. Amudha, "Review of Artificial Intelligent based Algorithm for Prediction of Polycystic Ovary Syndrome(PCOS) from Blood Samples," 2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2022, pp. 1172-1176, doi: 10.1109/ICIRCA54612.2022.9985699.

[16] Khanna, Varada Vivek, Krishnaraj Chadaga, Niranajana Sampathila, Srikanth Prabhu, Venkatesh Bhandage, and Govardhan K. Hegde. 2023. "A Distinctive Explainable Machine Learning Framework for Detection of Polycystic Ovary Syndrome" Applied System Innovation 6, no. 2: 32. https://doi.org/10.3390/asi6020032.

[17] Darji, M., Dave, J.A., Rathod, D.B. (2023). Review of Deep Learning: A New Era. In: Tuba, M., Akashe, S., Joshi, A. (eds) ICT Infrastructure and Computing. Lecture Notes in Networks and Systems, vol 520. Springer, Singapore. https://doi.org/10.1007/978-981-19-5331-6_33

[18] V. C. Gandhi and P. P. Gandhi, “A Survey-Insights of ML and DL in Health Domain,” in In2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), IEEE, 2022, pp. 239–246.

[19] Jain, A., Sharma, P.C., Vishwakarma, S.K., Gupta, N.K., Gandhi, V.C. (2022). Metaheuristic Techniques for Automated Cryptanalysis of Classical Transposition Cipher: A Review. In: Somani, A.K., Mundra, A., Doss, R., Bhattacharya, S. (eds) Smart Systems: Innovations in Computing. Smart Innovation, Systems and Technologies, vol 235. Springer, Singapore. https://doi.org/10.1007/978-981-16-2877-1_43

[20] A Vigorous Chaotic Function Based Image Authentication Structure AK Singh, VC Gandhi, MM Subramanyam, S Kumar, S Aggarwal, S Tiwari, Journal of Physics: Conference Series 1854 (1), 012039.