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

Machine Learning in Polycystic Ovary Syndrome (PCOS): A Comprehensive Review of Early Diagnosis, Personalized Treatment, and Predictive Insights.

Hauwa Ibrahim Aminu*[1], Moussa Mahamat Boukar*[2], Prema Kirubakaran[3]

Submitted:10/09/2024 **Revised:** 11/10/2024 **Accepted:** 20/10/2024

Abstract: Polycystic Ovary Syndrome (PCOS) is a multifaceted endocrine disorder that significantly affects women's health. This review examines the application of Machine Learning (ML) in PCOS research, focusing on its potential to enhance diagnosis, predict hormonal imbalances, and optimize personalized treatment plans. Current challenges in PCOS diagnosis and management, such as variability in diagnostic criteria and limited personalization of interventions, are outlined. The review systematically analyses recent advancements in ML techniques, highlighting their capabilities in addressing these challenges. Additionally, this paper identifies gaps in existing research, paving the way for future exploration of ML-driven innovations in PCOS management. By summarizing key findings, this review aims to provide a comprehensive understanding of the interplay between PCOS and ML while emphasizing the transformative potential of these technologies in improving patient outcomes.

Keywords: *Machine Learning, Polycystic Ovary Syndrome, Women, Hormones, Algorithm.*

1.0. Introduction

Polycystic Ovary Syndrome (PCOS) is a highly prevalent and complex endocrine disorder that affects 8–13% of women of reproductive age worldwide, with nearly 70% of cases remaining undiagnosed [1]. It manifests through a diverse set of symptoms, including irregular menstrual cycles, hyperandrogenism, ovarian cysts, and insulin resistance. Beyond its physical impact, PCOS has profound psychological consequences, contributing to anxiety, depression, and diminished quality of life [2]. The syndrome also significantly increases the risk of long-term health complications such as type 2 diabetes, cardiovascular disease, and infertility [3]. Despite decades of research, its etiology remains unclear, and diagnostic and treatment approaches are often inadequate or inconsistent [2] [3].

Machine Learning (ML), a subset of artificial intelligence, has recently emerged as a transformative tool in healthcare [4]. It enables the analysis of complex and multidimensional datasets, extracting patterns that would otherwise remain hidden. In PCOS research, ML offers unprecedented opportunities for early diagnosis, personalized treatment, and predicting long-term outcomes [1] [5]. By leveraging ML, researchers aim to bridge the gap between the growing availability of medical data and the need for actionable clinical insights. However, the integration of ML in PCOS research presents unique challenges, including a lack of standardized datasets, the inherent complexity of PCOS, and issues of interpretability in ML models.

Existing reviews and studies have explored various applications of ML in PCOS diagnosis and management [6]. However, many of these studies focus narrowly on specific ML techniques, leaving significant gaps in our understanding of their broader implications for clinical practice. Furthermore, while PCOS research spans numerous dimensions, from genetics and hormonal regulation to lifestyle interventions, few studies synthesize these aspects in the context of ML applications [7].

This review aims to provide a comprehensive analysis of ML applications in PCOS research, addressing gaps in existing literature. It explores ML's potential to enhance diagnosis, optimize personalized management, and predict disease progression, while also evaluating limitations in current approaches. By critically analysing recent studies, this paper seeks to illuminate trends, identify challenges, and propose future research directions to advance PCOS research through ML integration.

The review is organized as follows: Section 2 presents the background of PCOS, highlighting its symptoms, complications, and current challenges, alongside an introduction to ML and its relevance to healthcare. Section 3 reviews related works to provide context for this study. Section 4 outlines the systematic methodology adopted for this review. Section 5 discusses key findings, trends, and challenges. Finally, Section 6 concludes with recommendations for future research.

2.0. Background

Polycystic Ovary Syndrome (PCOS) is a complex hormonal disorder that affects millions of women worldwide, primarily those

1 Hauwa Ibrahim Aminu, Nile University of Nigeria, Abuja – NIGERIA
ORCID ID : 0009-0005-1033-5183
2 Moussa Mahamat Boukar, Nile University of Nigeria, Abuja – NIGERIA
ORCID ID : 0000-3343-7165-777X
3 Prema Kirubakaran, Nile University of Nigeria, Abuja – NIGERIA
ORCID ID : 0000-3343-7165-777X
** Corresponding Author Email: hauwa.aminu@outlook.com*

of reproductive age. Characterized by diverse symptoms and significant variation in presentation among individuals, PCOS presents unique challenges in diagnosis and management. The disorder is associated with irregular menstrual cycles, hyperandrogenism (excess androgen levels), polycystic ovaries (see figure 1) detectable via ultrasound, and insulin resistance [8], [9]. These symptoms can contribute to a wide range of complications, including infertility, obesity, type 2 diabetes, cardiovascular risks, and psychological distress such as anxiety and depression. Despite its prevalence, affecting approximately 8-13% of women globally, the underlying causes of PCOS remain poorly understood, and up to 70% of cases go undiagnosed [2], [3].

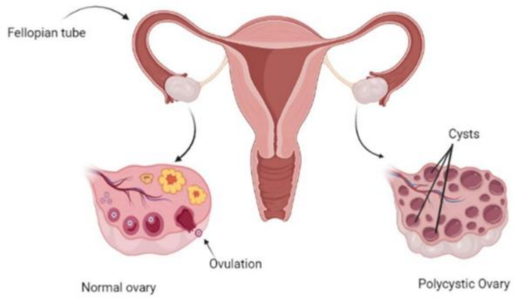

Figure 1 Visual Representation of PCOS [10]

2.1. Overview of PCOS Challenges

The diagnostic process for PCOS is particularly challenging due to its heterogeneous nature. While criteria such as the Rotterdam criteria provide a standardized approach, relying on patient-reported symptoms and subjective evaluations often leads to misdiagnosis or delayed diagnosis [9]. Furthermore, PCOS has broader implications, including long-term risks for metabolic syndrome, endometrial cancer, and other complications that require timely intervention [11], [12].

2.2. Advancements and Gaps in PCOS Research

Research into PCOS has evolved significantly since its early identification as Stein-Leventhal Syndrome in the 1930s [13], [14]. Advances in imaging techniques and genomic studies have enhanced our understanding of its clinical and genetic underpinnings [15]. For instance, genome-wide association studies (GWAS) have identified genetic markers linked to PCOS susceptibility, while improved imaging technologies have facilitated the detailed visualization of ovarian morphology [6], [10]. However, these advancements have not been universally accessible or applied, creating disparities in PCOS diagnosis and management across populations.

2.3. Machine Learning in PCOS Research

The application of machine learning (ML) in PCOS research offers transformative potential, addressing many of the challenges in diagnosis, prediction, and treatment [16]. ML algorithms can analyse vast datasets—including clinical, hormonal, and genetic data—uncovering patterns and relationships that may otherwise remain obscured [15]. For example, ML techniques such as supervised learning have been applied to predict PCOS risk factors, while unsupervised learning models cluster patients based on symptomatology for more personalized care [17], [18].

In an era where there is rapid integration of advanced technology in various segments, specifically in the healthcare arena, the synergy that lies between machine learning and the complexities associated with PCOS offers a compelling frontier of research and innovation [19]. While PCOS remains obscure, the enigma it lies in provides an opportunity for advancement. In the world of healthcare, cutting-edge technology such as machine learning has been used to aid in certain diagnosis for various diseases including PCOS. With machine learning, it is possible to explore the intricate landscape of PCOS and in doing so, enhancing diagnosis and having a more general understanding of the syndrome [20], [21]. As the scientific field continues to advance in terms of deciphering the intricacies of PCOS, so does the tech field [22]. ML has emerged as a powerful tool which has so far brough about numerous positive additions such as personalized treatment, risk prediction and diagnosis in general, with regards to PCOS [23].

The integration of machine learning in medicine has transformed healthcare practices. From automating medical image interpretation for diagnoses [24], [25] to supporting clinical decision-making [26], drug discovery [27], and even enabling health monitoring via wearables [28]. These tools hold the promise of improving early detection, optimizing treatment pathways, and enhancing patient outcomes.

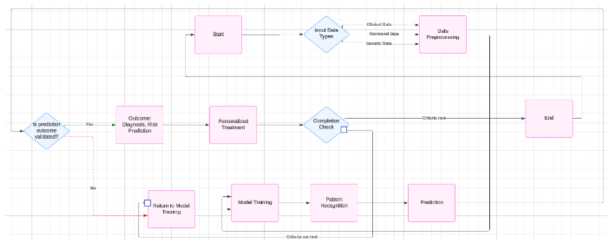

Figure 2 Workflow of Machine Learning in PCOS Research. This flowchart illustrates how ML processes—from data collection to clinical application—address challenges in PCOS diagnosis, treatment, and risk prediction.

2.4. Bridging the Gaps: The Role of Systematic Reviews

Given the complexity of PCOS and the emerging role of ML, there is a pressing need for systematic reviews to synthesize existing research. These reviews can identify trends, gaps, and opportunities for interdisciplinary collaboration, providing a roadmap for advancing PCOS research. By examining the intersection of PCOS and ML, this paper seeks to illuminate how computational tools can enhance our understanding and management of this enigmatic syndrome.

2.5. Limitations of Current Approaches

Despite the promise of ML, its integration into PCOS research and clinical practice remains in its infancy. Existing studies often lack diverse datasets that encompass variations in populations, phenotypes, and geographic locations [29]. Furthermore, many ML models are developed and tested in isolation without consideration for interdisciplinary approaches that combine clinical expertise with advanced computational methods. Ethical concerns surrounding data privacy and the interpretability of complex ML models also pose barriers to widespread adoption. [30]

3.0. Materials and Methodology

The methodology for this review adheres to a systematic review framework, ensuring a comprehensive and unbiased synthesis of existing literature on machine learning (ML) applications in Polycystic Ovary Syndrome (PCOS). This approach was structured to identify, evaluate, and integrate findings from relevant studies to address the primary objectives of this review.

The overall workflow for the systematic review process is illustrated in the figure below. This diagram outlines the major steps taken, from defining objectives to the final thematic analysis.

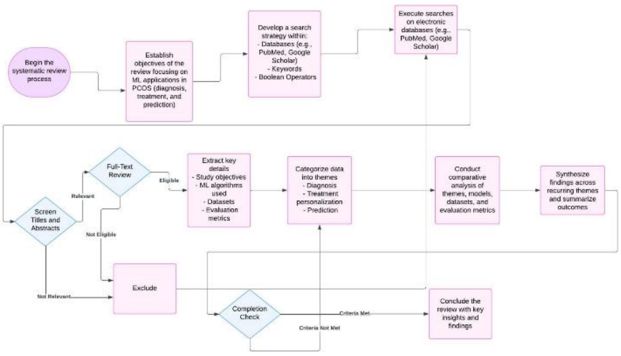

Figure 3 Flowchart of the systematic methodology for reviewing ML applications in PCOS.

The following subsections provide detailed descriptions of each stage in the process.

3.1. Search Strategy

The search process utilized a combination of electronic databases, including PubMed, Google Scholar, and specialized repositories focused on PCOS and machine learning research. Keywords and search phrases were carefully designed to capture a wide range of studies, such as:

- **Keywords**: "Polycystic Ovary Syndrome," "PCOS," "Machine Learning," "ML in PCOS," "Diagnosis of PCOS using ML," "Predictive Models for PCOS," and "Treatment Personalization in PCOS."

- **Boolean Operators**: Queries were refined using Boolean operators (AND, OR) to combine relevant terms (e.g., "PCOS AND Machine Learning" OR "Prediction Models for PCOS").

The search period was limited to studies published within the last five years (2019–2024) to ensure the inclusion of up-to-date research while allowing for foundational insights.

3.2. Inclusion and Exclusion Criteria

Studies were selected based on predefined inclusion and exclusion criteria to maintain focus and relevance:

- **Inclusion Criteria**:

1. Studies applying ML algorithms to PCOS research, focusing on diagnosis, prediction, or treatment personalization.

2. Peer-reviewed articles and conference proceedings.

3. Studies providing sufficient methodological detail, including datasets, algorithms, and evaluation metrics.

- **Exclusion Criteria**:

1. Studies without clear applications of ML in PCOS.

2. Publications lacking methodological transparency or comprehensive results.

3. Articles published in non-English languages (due to translation limitations).

3.3. Data Extraction and Categorization

Following the selection process, data from included studies were systematically extracted and categorized. Key details recorded from each study included:

1. **Study Objectives**: The primary focus and goals of each research effort.

2. **ML Algorithms Used**: Specific models (Logistic Regression, SVM, Random Forest, Linear Regression) and their applications.

3. **Datasets**: Characteristics of datasets, including size, type (e.g., clinical, hormonal, genetic), and diversity.

4. **Evaluation Metrics**: Methods used to assess model performance, such as accuracy, precision, recall, and AUC-ROC.

These data points were organized into a structured matrix, facilitating comparative analysis and thematic grouping.

3.4. Focus Areas

To provide a comprehensive review, the analysis was categorized into three key focus areas: diagnosis, treatment personalization, and prediction using ML.

3.4.1. Diagnosis of PCOS

Studies under this theme explored the use of ML algorithms to enhance diagnostic accuracy by analyzing clinical, hormonal, and imaging data. Emphasis was placed on:

- Supervised Learning Models: Techniques such as support vector machines (SVMs) and random forests were evaluated for their ability to classify PCOS patients based on diagnostic criteria [31].

- Deep Learning Approaches: Neural networks, particularly convolutional neural networks (CNNs), demonstrated effectiveness in analyzing ultrasound images and identifying ovarian cysts [32].

3.4.2. Treatment Personalization

This segment focuses on ML-driven approaches to customize treatment plans based on individual patient characteristics. Key findings included:

- The use of ensemble methods to predict responses to lifestyle interventions, such as diet and exercise [33].

- Algorithms capable of optimizing pharmacological treatments by analyzing hormonal and metabolic profiles [34].

3.4.3. Prediction of Outcomes

Under this focus area, studies applied predictive modelling to forecast PCOS-related risks, such as infertility, obesity, and diabetes. Commonly employed models included:

- Linear Regression: Predicting hormonal imbalances and their impact on long-term health [35].

- Reinforcement Learning: Exploring adaptive strategies for dynamic management of symptoms [36].

3.5. Key Themes in Analysis

The review synthesized findings across several recurring themes:

1. Algorithms: A comparison of commonly used ML models, highlighting their strengths and limitations in PCOS applications.

2. Datasets: Analysis of data diversity, including genetic, clinical, and demographic factors, and their impact on model performance.

3. Evaluation Metrics: Discussion of how metrics like accuracy and AUC-ROC were used to evaluate and compare model efficacy.

4.0. Findings and Discussion

This section synthesizes key findings from the reviewed literature, categorizing them into three major areas: diagnosis, treatment personalization, and prediction of outcomes. It highlights the specific contributions of various studies, the methodologies employed, and the results achieved, offering a comprehensive review of how ML has advanced PCOS research.

4.1. Diagnosis

ML techniques have shown significant potential to improve the accuracy and reliability of PCOS diagnosis by analyzing diverse datasets and extracting patterns beyond human interpretation.

4.1.1. Supervised Learning Models:

Studies such as [23] employed Random Forest in their study "Predicting polycystic ovary syndrome (PCOS) with machine learning algorithms from electronic health records" to predict PCOS prior to clinical diagnosis. They attained this by classifying patients based on clinical and hormonal features, attaining an 82% accuracy. These models emphasize the importance of features such as insulin resistance and body mass index (BMI).

Logistic Regression, as demonstrated by [20] in their paper titled "Polycystic Ovary Syndrome Detection Machine Learning Model Based on Optimized Feature Selection and Explainable Artificial Intelligence", has been particularly effective for binary classification tasks, offering an interpretable framework for clinicians. The main objective of this paper is to assist in early detection of PCOS diagnosis while at the same time aiming to reduce the problems and any complications caused by the syndrome. They achieved 90% accuracy. Figure 4 showcases the performance evaluation for some of the ML models used in the experiment [20].

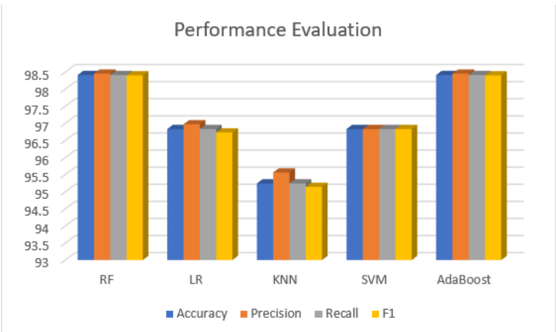

Figure 4 Shows the Performance Evaluation for Random Forest, Logistic Regression, k-Nearest Neighbors and AdaBoost Models. This is done by using the Recursive Feature Elimination Feature Selection Method [20]

KNN is being leveraged in PCOS research for clustering and classifying individuals based on similar features. KNN offers a more flexible and intuitive approach with regards to PCOS data. This makes it very beneficial when it comes to understanding the complexity of the syndrome and to enhance diagnosis. For instance, [33] in their research "Exploring the dominant features and data-driven detection of polycystic ovary syndrome through modified stacking ensemble machine learning technique", utilized KNN. In this study, they identified the subgroups within PCOS population based on clinical parameters [33]. The table and graph below showcase a summary of the performance evaluation for some of the models used in the study:

Table 1. Showing performance evaluation of models [33]

	Performance Evaluation			
Model	***Accuracy (in %)***	***Precision (in %)***	***Recall (in %)***	***F1 (in %)***
SMV	89	87.4	88.6	88
Log. Reg	89	87.8	87.8	87.8
Decision Tree	82.2	80.2	80.5	80.4
KNN	81.6	79.8	82.2	80.5
NaiveBayes	84	84.1	79.8	81.2
RandomForest	89	87.8	87.8	87.8
AdaBoosting	89.6	88.6	88.2	88.4

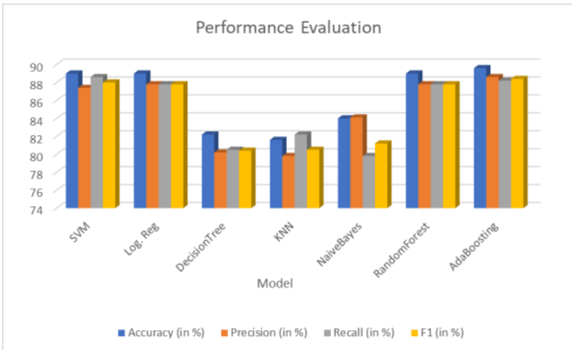

Figure 5 Comparing Models Performance [33]

4.1.2. Deep Learning Techniques:

Convolutional Neural Networks (CNNs) were utilized by [32] in the study titled "Deep learning approach for ovarian cysts detection and classification (OCD-FCNN) using fuzzy convolutional neural network" to analyze ovarian ultrasound images, achieving a diagnostic precision of 98.37% which outperformed traditional methods reliant on visual assessment. This approach reduces subjectivity in diagnosis and provides automated detection capabilities [32].

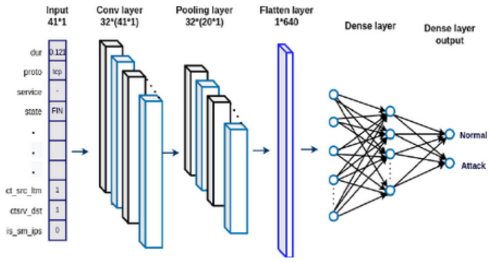

Figure 6 CNN Architecture designed to enhance the performance of network intrusion detection systems (IDS) by effectively extracting and learning high-level features from network traffic data [49]

4.2. Treatment Personalization

One of the most promising applications of ML in PCOS research is the development of predictive models for personalized treatment plans.

4.2.1. Predictive Models:

[31] applied Support Vector Machines (SVMs) as part of a hybrid algorithm to predict presence of PCOS. The PSO-SVM algorithm achieved a maximum accuracy of 88.96% in predicting PCOS using a Linear SVM kernel. The study aimed to improve PCOS diagnosis for gynecologists, leading to faster and more accurate treatment plans, while empowering patients to take early preventive actions and simplifying the diagnostic process.

Ensemble methods like AdaBoost, as used by [33], have been applied in PCOS research for improving predictive accuracy and ensemble learning.[33] used AdaBoost to explore and enhance the dominant features of PCOS by amalgamating the predictive power of multiple weak models to achieve vigorous results [33].

4.2.2. Feature Importance Analysis:

Random Forest models, implemented in studies like [23] titled “Predicting polycystic ovary syndrome (PCOS) with machine learning algorithms from electronic health records”, identified insulin resistance, androgen levels, and BMI as critical predictors of treatment outcomes. These insights provide clinicians with actionable parameters for tailoring treatment.

4.3. Prediction of Outcomes and Risk Factors

ML has also been extensively applied to forecast long-term risks and complications associated with PCOS, such as infertility, diabetes, and cardiovascular disease.

4.3.1. Longitudinal Analysis:

Using Long Short-Term Memory Networks (LSTMs), [37] modelled temporal patterns in hormonal fluctuations and menstrual irregularities. The study used LSTM networks to predict menstrual cycle lengths from time series data. The LSTM model achieved strong performance metrics: MAE of 0.3000, MSE of 0.5477, and RMSE of 0.7401, indicating its effectiveness compared to other models, although no specific accuracy percentage was provided.

4.3.2. Linear Regression:

The use of Linear Regression models for PCOS research has also been noted, particularly in predicting hormonal imbalances and metabolic challenges [38]. These models provide interpretable results, enabling clinicians to make initial assessments based on features such as body weight, hormonal levels, and clinical history. For instance, one study utilized Linear Regression to predict

Table 1. Comparative Study of ML Algorithms in PCOS Research

Algorithms	*Authors (Year)*	*Use Case*	*Dataset Type*	*Accuracy (%)*	*Strengths*	*Limitations*
Random Forest	(Zad et al., 2023)	Diagnosis of PCOS	Electronic Health Records	82	Handles mixed data; identifies critical features like BMI and insulin resistance	Computationally intensive
Logistic Regression	(Elmannai et al.,2023)	Early Detection of PCOS	Clinical Data	90	Interpretable; effective for binary classification tasks	Limited to linear relationships
KNN	(Suha&Islam, 2023)	Clustering PCOS Symptoms	Clinical Parameters	87.8	Intuitive and flexible for clustering	Sensitive to irrelevant features
SVM	(Jantan, Fatihah, Bahrin, & Shaufee, 2024)	Diagnosis and Treatment Planning	Clinical Data	88.86	Effective with small datasets; robust classification	High computational cost
CNN (Deep Learning)	(Jantan et al., 2024)	Ovarian Ultrasound Image Analysis	Ultrasound Image	98.37	Excels in image analysis; automates diagnosis	Requires large datasets
AdaBoost (Ensemble Method)	(Suha&Islam, 2023)	Feature Selection and PCOS Prediction	Clinical Data	95.7	Combines weak learners for strong results	Requires careful parameter tuning
Long Short-Term Memory (LSTM)	(Rego, 2023)	Predicting Hormonal Fluctuations	Time-Series Data	MAE: 0.3000, MSE: 0.5477, RMSE: 0.7401	Effective in modeling temporal patterns	No specific accuracy percentage provided
Linear Regression	(H & Anusuya, 2020)	Predicting Hormonal Imbalances	Clinical Data	N/A	Simple; interpretable for initial assessments	Limited for complex, non-linear problems

testosterone levels as a marker for hyperandrogenism, a critical diagnostic feature of PCOS [35].

4.3.3. Risk Stratification:

Studies such as [23] employed Random Forest models to stratify patients into high-risk and low-risk groups by incorporating genetic, clinical, and lifestyle data. Their approach demonstrated an area under the curve (AUC) of 0.88, underscoring the discriminatory ability of these methods.

4.4. Trends and Insights

Several trends emerged from the reviewed studies, reflecting advancements and challenges in the application of ML to PCOS research:

4.4.1. Common ML Techniques:

Supervised Learning: Logistic Regression and Random Forest dominate classification tasks due to their interpretability and scalability [20] [39].

Deep Learning: Techniques such as CNNs and LSTMs are increasingly favoured for tasks involving complex pattern recognition, such as image analysis and time-series forecasting [32], [37]. To provide a clearer understanding of the application of machine learning algorithms in PCOS research, a comparative analysis of their use cases, dataset types, performance metrics, strengths, and limitations is presented in Table 2.

4.4.2. Datasets:

Most studies relied on electronic health records, genetic data, or imaging datasets, highlighting the need for multimodal approaches to capture the complexity of PCOS. Challenges remain in achieving data diversity and standardization across populations [40].

4.4.3. Strengths and Limitations of ML Approaches:

ML models have demonstrated their ability to integrate diverse datasets, providing insights into diagnosis, treatment, and outcomes [40]. However, the reliance on small, localized datasets limits the generalizability of findings to broader populations. Also, deep learning models, while accurate, face challenges related to interpretability [29], [41].

4.5. Challenges

Despite promising advancements, several challenges hinder the effective integration of ML in PCOS research. Firstly, there is lack of standardized datasets [29]. The absence of standardized, diverse datasets limits cross-study comparisons and model validation. Variability in data sources, including differences in clinical, genetic, and demographic characteristics, hinders reproducibility. Secondly, there are issues with model interpretability [42]. Complex models such CNNs and LSTMs, while accurate, often lack transparency, making them difficult for clinicians to interpret and trust in a clinical setting [32], [37]. There are also ethical considerations to take into account. Privacy concerns regarding the use of sensitive data, such as genetic information, remain significant [43]. Data biases due to underrepresentation of certain populations lead to disparities in outcomes [44].

4.6. Conclusion of Findings

The reviewed studies underscore the transformative potential of ML in PCOS research, from enhancing diagnostic accuracy to enabling personalized treatments. However, significant efforts are required to overcome challenges related to data standardization, interpretability, and ethical concerns. Future research should focus on creating diverse, standardized datasets and developing interpretable models to bridge the gap between computational advancements and clinical implementation.

5.0. Future Directions

The integration of machine learning (ML) into Polycystic Ovary Syndrome (PCOS) research is a promising frontier. However, several critical gaps must be addressed to realize its full potential. Based on the limitations highlighted in this review, the following directions are proposed:

5.1. Research Gaps

There is need for a diverse and standardized dataset. In these studies, there is a lack of universally accepted datasets limits cross-study comparisons and the validation of ML models. Current datasets often fail to represent the heterogeneity of PCOS manifestations across populations [45]. Standardized and diverse datasets encompassing demographic, genetic, clinical, and lifestyle factors are urgently needed [46]. Therefore, future research must prioritize creating open-access repositories and initiatives to ensure inclusivity and minimize biases.

Another research gap is the integration of ML models with clinical workflows. Despite advancements in ML, there is a disconnect between computational models and their adoption in clinical settings. Bridging this gap requires developing user-friendly, interpretable tools that align with clinicians' decision-making processes [29].

5.2. Proposed Solutions

It is important to have interdisciplinary collaborations. Successful integration of ML into PCOS research demands collaboration among clinicians, data scientists, geneticists, and ethicists. These collaborations can ensure that models are clinically relevant, ethically sound, and technically robust. For example, explainable AI techniques could be employed to make model outputs transparent and actionable.

Another proposed solution is addressing ethical and privacy concerns. Ethical issues, such as data privacy and the potential for algorithmic bias, remain critical challenges. ML models must comply with stringent data protection regulations, such as GDPR, and incorporate fairness metrics to avoid perpetuating disparities in healthcare outcomes. A focus on ethical AI is essential for fostering trust among stakeholders.

5.3. Emerging Technologies

Deep Learning and Explainable AI are two emerging technologies that can be further explored. Advanced deep learning models, including Convolutional Neural Networks (CNNs) and Long Short-Term Memory Networks (LSTMs), have shown promise in analyzing complex datasets [47]. However, the adoption of explainable AI techniques is critical to improve interpretability and facilitate clinical integration [48].

Additionally, leveraging big data frameworks can enable the integration of multimodal datasets, such as genetic profiles, imaging data, and electronic health records (EHRs). This holistic approach can uncover novel insights into PCOS while enhancing the scalability of ML applications [23].

6.0. Conclusion

Ultimately, this review has thoroughly examined the intersection of PCOS and machine learning, while highlighting the critical aspects and advancements. Additionally, it addresses the key challenges associated with Polycystic Ovary Syndrome (PCOS). By synthesizing findings across diagnosis, treatment personalization, and prediction of outcomes, this study underscores ML's ability to revolutionize PCOS research and clinical management.

In this review, some of the key findings include machine learning models being used for PCOS prediction and why in addition to other result. Machine learning models, such as Random Forest, Logistic Regression, and CNNs, demonstrate high accuracy in diagnosing PCOS and predicting outcomes. Personalized treatment approaches, driven by predictive modelling, show promise in optimizing interventions and improving patient outcomes. Despite these advancements, challenges remain, including data standardization, model interpretability, and ethical concerns.

This review addresses gaps in the existing literature by providing a comprehensive analysis of ML applications in PCOS research. It highlights trends, challenges, and emerging opportunities, offering a roadmap for future research efforts.

To unlock the full potential of ML in PCOS research, interdisciplinary collaboration is paramount. Clinicians, data scientists, and policymakers must work together to develop diverse, standardized datasets, interpretability-focused models, and ethically responsible AI tools. Bridging computational advancements and clinical workflows will be critical in translating ML research into impactful healthcare solutions.

In conclusion, the future of PCOS research lies at the intersection of advanced technology and human-centred approaches. By addressing existing challenges and embracing collaborative innovation, ML has the potential to transform the diagnosis, treatment, and long-term management of PCOS, ultimately improving outcomes for millions of women worldwide.

7.0. References

[1] S. A. Suha and M. N. Islam, "An extended machine learning technique for polycystic ovary syndrome detection using ovary ultrasound image," Sci Rep, vol. 12, no. 1, pp. 1–16, Dec. 2022, doi: 10.1038/s41598-022-21724-0.

[2] S. Singh et al., "Polycystic Ovary Syndrome: Etiology, Current Management, and Future Therapeutics," J Clin Med, vol. 12, no. 4, p. 12, Feb. 2023, doi: 10.3390/JCM12041454.

[3] S. Zhuang, C. Jing, L. Yu, L. Ji, W. Liu, and X. Hu, "The relationship between polycystic ovary syndrome and infertility: a bibliometric analysis," Ann Transl Med, vol. 10, no. 6, pp. 318–318, Mar. 2022, doi: 10.21037/ATM-22-714.

[4] N. M. Bogari, "Genetic construction between polycystic ovarian syndrome and type 2 diabetes," Saudi J Biol Sci, vol. 27, no. 10, pp. 2539–2543, Oct. 2020, doi: 10.1016/J.SJBS.2020.05.004.

[5] M. J. Javid-Naderi, A. Mahmoudi, P. Kesharwani, T. Jamialahmadi, and A. Sahebkar, "Recent advances of nanotechnology in the treatment and diagnosis of polycystic ovary syndrome," J Drug Deliv Sci Technol, vol. 79, p. 104014, Jan. 2023, doi: 10.1016/J.JDDST.2022.104014.

[6] S. Akre, K. Sharma, S. Chakole, and M. B. Wanjari, "Recent Advances in the Management of Polycystic Ovary Syndrome: A Review Article", doi: 10.7759/cureus.27689.

[7] S. B. Christensen et al., "The prevalence of polycystic ovary syndrome in adolescents," Fertil Steril, vol. 100, no. 2, pp. 470–477, Aug. 2013, doi: 10.1016/J.FERTNSTERT.2013.04.001.

[8] Y. V. Louwers and J. S. E. Laven, "Characteristics of polycystic ovary syndrome throughout life," https://doi.org/10.1177/2633494120911038, vol. 14, p. 263349412091103, Mar. 2020, doi: 10.1177/2633494120911038.

[9] J. Bulsara, P. Patel, A. Soni, and S. Acharya, "A review: Brief insight into Polycystic Ovarian syndrome," Endocrine and Metabolic Science, vol. 3, 2021, doi: 10.1016/j.endmts.2021.100085.

[10] H. Nautiyal et al., "Polycystic Ovarian Syndrome: A Complex Disease with a Genetics Approach," Biomedicines, vol. 10, no. 3, Mar. 2022, doi: 10.3390/BIOMEDICINES10030540.

[11] J. V. Pinkerton, "Polycystic Ovary Syndrome (PCOS) - ." Accessed: Sep. 21, 2023. [Online]. Available: https://www.msdmanuals.com/professional/gynecology-and-obstetrics/menstrual-abnormalities/polycystic-ovary-syndrome-pcos#

[12] A. M. Rababa'h, B. R. Matani, and A. Yehya, "An update of polycystic ovary syndrome: causes and therapeutics options," Heliyon, vol. 8, no. 10, p. e11010, Oct. 2022, doi: 10.1016/J.HELIYON.2022.E11010.

[13] J. A. Lentscher and A. H. Decherney, "Clinical Presentation and Diagnosis of Polycystic Ovarian Syndrome," Clin Obstet Gynecol, vol. 64, no. 1, 2020, Accessed: Sep. 21, 2023. [Online]. Available: www.clinicalobgyn.com

[14] A. Agrawal, R. Ambad, R. Lahoti, P. Muley, and P. Pande, "Role of artificial intelligence in PCOS detection," Journal of Datta Meghe Institute of Medical Sciences University, vol. 17, no. 2, p. 491, Apr. 2022, doi: 10.4103/JDMIMSU.JDMIMSU_278_22.

[15] B. Mahesh, "Machine Learning Algorithms -A Review." Accessed: Sep. 21, 2023. [Online]. Available: https://www.researchgate.net/publication/344717762_Machine_Learning_Algorithms_-A_Review

[16] T. J. Cleophas and A. H. Zwinderman, "Machine learning in medicine - a complete overview," Machine Learning in Medicine - A Complete Overview, pp. 1–667, Mar. 2020, doi: 10.1007/978-3-030-33970-8/COVER.

[17] S. Sah, "Machine Learning: A Review of Learning Types," Jul. 2020, doi: 10.20944/PREPRINTS202007.0230.V1.

[18] A. Usman, M. M. Boukar, M. A. Suleiman, and I. A. Salihu, "Test Case Generation Approach for Android Applications using Reinforcement Learning," Engineering, Technology & Applied Science Research, vol. 14, no. 4, pp. 15127–15132, Aug. 2024, doi: 10.48084/ETASR.7422.

[19] S. Bharati, P. Podder, and M. R. Hossain Mondal, "Diagnosis of Polycystic Ovary Syndrome Using Machine Learning Algorithms," 2020 IEEE Region 10 Symposium, TENSYMP 2020, pp. 1486–1489, Jun. 2020, doi: 10.1109/TENSYMP50017.2020.9230932.

[20] H. Elmannai et al., "Polycystic Ovary Syndrome Detection Machine Learning Model Based on Optimized Feature Selection and Explainable Artificial Intelligence," Diagnostics, vol. 13, no. 8, Apr. 2023, doi: 10.3390/DIAGNOSTICS13081506.

[21] 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," IEEE Region 10

Annual International Conference, Proceedings/TENCON, pp. 673–678, Oct. 2019, doi: 10.1109/TENCON.2019.8929674.

[22] F. J. Barrera et al., "Application of machine learning and artificial intelligence in the diagnosis and classification of polycystic ovarian syndrome: a systematic review," Front Endocrinol (Lausanne), vol. 14, p. 1106625, Sep. 2023, doi: 10.3389/FENDO.2023.1106625.

[23] Z. Zad et al., "Predicting polycystic ovary syndrome (PCOS) with machine learning algorithms from electronic health records," medRxiv, p. 2023.07.27.23293255, Aug. 2023, doi: 10.1101/2023.07.27.23293255.

[24] A. Barragán-Montero et al., "Artificial intelligence and machine learning for medical imaging: a technology review HHS Public Access," Phys Med, vol. 83, pp. 242–256, Mar. 2021, doi: 10.1016/j.ejmp.2021.04.016.

[25] H. Alkemade, "Medical Imaging with Azure Machine Learning ." Accessed: Sep. 19, 2023. [Online]. Available: https://towardsdatascience.com/medical-imaging-with-azure-machine-learning-b5acfd772dd5

[26] M. R. Afrash et al., "Machine Learning-Based Clinical Decision Support System for Automatic Diagnosis of COVID-19 based on Clinical Data," Journal of Biostatistics and Epidemiology, vol. 8, no. 1, pp. 77–89, 2022, doi: 10.18502/JBE.V8I1.10407.

[27] G. Jain, "Application of Machine Learning in Drug Discovery and Development Lifecycle," International Journal of Medical, Pharmacy and Drug Research, vol. 6, no. 6, pp. 16–35, 2022, doi: 10.22161/ijmpd.6.6.4.

[28] K. P. Seng, L.-M. Ang, E. Peter, and A. Mmonyi, "Machine Learning and AI Technologies for Smart Wearables," Electronics (Basel), vol. 12, no. 7, p. 1509, Mar. 2023, doi: 10.3390/electronics12071509.

[29] R. Kenge, "Machine Learning, Its Limitations, and Solutions Over IT," International Journal of Applied Research on Information Technology and Computing, vol. 11, no. 2, p. 73, 2020, doi: 10.5958/0975-8089.2020.00009.3.

[30] G. Varoquaux and V. Cheplygina, "Machine learning for medical imaging: methodological failures and recommendations for the future," NPJ Digit Med, vol. 5, no. 1, p. 48, Apr. 2022, doi: 10.1038/s41746-022-00592-y.

[31] H. Jantan, U. Fatihah, M. Bahrin, and L. H. Shaufee, "Polycystic Ovary Syndrome (PCOS) Prediction System Using PSO-SVM," Journal of Computing Research and Innovation, vol. 9, no. 1, p. 2024, 2024, doi: 10.24191/jcrinn.v9i1.

[32] T. Nadana Ravishankar, H. Makarand Jadhav, N. S. Kumar, S. Ambala, and M. Pillai, "A deep learning approach for ovarian cysts detection and classification (OCD-FCNN) using fuzzy convolutional neural network," Measurement: Sensors, vol. 27, p. 100797, 2023, doi: 10.1016/j.measen.2023.100797.

[33] S. Alam 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, p. e14518, Mar. 2023, doi: 10.1016/j.heliyon.2023.e14518.

[34] M. Szczuko et al., "Nutrition Strategy and Life Style in Polycystic Ovary Syndrome—Narrative Review," Nutrients, vol. 13, no. 7, Jul. 2021, doi: 10.3390/NU13072452.

[35] P. H and M. A. Anusuya, "A Prediction Model for Evaluating the Risk of Developing PCOS," International Research Journal of Engineering and Technology, 2020, Accessed: Sep. 27, 2023. [Online]. Available: www.irjet.net

[36] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," May 01, 2021, Springer. doi: 10.1007/s42979-021-00592-x.

[37] R. Rego, "Predictive Modeling of Menstrual Cycle Length: A Time Series Forecasting Approach," Jun. 2023.

[38] Y. Guo et al., "A structural equation model linking health literacy, self-efficacy, and quality of life in patients with polycystic ovary syndrome," BMC Womens Health, vol. 23, no. 1, pp. 1–9, Dec. 2023, doi: 10.1186/S12905-023-02223-4/FIGURES/1.

[39] Z. Zad et al., "Predicting polycystic ovary syndrome (PCOS) with machine learning algorithms from electronic health records," medRxiv, Aug. 2023, doi: 10.1101/2023.07.27.23293255.

[40] F. J. Barrera et al., "Application of machine learning and artificial intelligence in the diagnosis and classification of polycystic ovarian syndrome: a systematic review," Front Endocrinol (Lausanne), vol. 14, 2023, doi: 10.3389/FENDO.2023.1106625/FULL.

[41] S. Chang and A. Dunaif, "Diagnosis of Polycystic Ovary Syndrome: Which Criteria to Use When?," Endocrinol Metab Clin North Am, vol. 50, no. 1, p. 11, Mar. 2021, doi: 10.1016/J.ECL.2020.10.002.

[42] R. Kaur, R. Kumar, and M. Gupta, "Food Image-based diet recommendation framework to overcome PCOS problem in women using deep convolutional neural network," Computers and Electrical Engineering, vol. 103, p. 108298, Oct. 2022, doi: 10.1016/J.COMPELECENG.2022.108298.

[43] R. Tiwari, "Ethical And Societal Implications of AI and Machine Learning," 2023, doi: 10.55041/IJSREM17505.

[44] L. Seyyed-Kalantari, H. Zhang, M. B. A. McDermott, I. Y. Chen, and M. Ghassemi, "Reply to: 'Potential sources of dataset bias complicate investigation of underdiagnosis by machine learning algorithms' and 'Confounding factors need to be accounted for in assessing bias by machine learning algorithms,'" Nature Medicine 2022 28:6, vol. 28, no. 6, pp. 1161–1162, Jun. 2022, doi: 10.1038/S41591-022-01854-8.

[45] S. Cowan et al., "Lifestyle management in polycystic ovary syndrome – beyond diet and physical activity," BMC Endocr Disord, vol. 23, no. 1, Dec. 2023, doi: 10.1186/S12902-022-01208-Y.

[46] J. Liu et al., "Measuring the global disease burden of polycystic ovary syndrome in 194 countries: Global Burden of Disease Study 2017," Human Reproduction, vol. 36, no. 4, pp. 1108–1119, Mar. 2021, doi: 10.1093/humrep/deaa371.

[47] R. DiPietro and G. D. Hager, "Deep learning: RNNs and LSTM," Handbook of Medical Image Computing and Computer Assisted Intervention, pp. 503–519, Jan. 2019, doi: 10.1016/B978-0-12-816176-0.00026-0.

[48] M. H. Kangasniemi et al., "Artificial intelligence deep learning model assessment of leukocyte counts and proliferation in endometrium from women with and without polycystic ovary syndrome," F and S Science, vol. 3, no. 2, pp. 174–186, May 2022, doi: 10.1016/j.xfss.2022.01.006.

[49] A. Meliboev, J. Alikhanov, and W. Kim, "Performance Evaluation of Deep Learning Based Network Intrusion Detection System across Multiple Balanced and Imbalanced Datasets," Electronics (Switzerland), vol. 11, no. 4, Feb. 2022, doi: 10.3390/ELECTRONICS11040515.