iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.51615

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

Comparative Study on AI Technology in PCOS Diagnosis

Aswathy V S[1] Vineetha Sankar P[2]
[1, 2]*Assistant Professor, Department of Computer Science, St. Albert's College (Autonomous), India*

Abstract: ***With the help of AI, healthcare systems can process massive volumes of clinical data with pinpoint accuracy. Polycystic ovarian syndrome (PCOS) is a disorder that affects women of reproductive age and causes hormonal imbalances, and it is a prevalent health issue among young women. A hormonal imbalance is the root cause of menstrual irregularities. Women with polycystic ovary syndrome are more likely to experience significant weight gain, increased facial hair, acne, hair loss, skin tone changes, and irregular menstruation, which in rare cases can result in infertility. Since a correct diagnosis is vital to effective treatment, this essay will examine and contrast several machine learning methods. We will also talk about an artificial intelligence method that combines aspects of both heterogeneous ML and Deep Learning.***
Keywords: Polycystic Ovarian Syndrome, Machine Learning, Random Forest, Deep learning, Artificial Intelligence

I. INTRODUCTION

Artificial intelligence (AI) is an all-encompassing technology that encourages individuals to rethink their information-gathering, data-analysis, and decision-making processes. Already, it affects every facet of people's lives. Machine learning is a subfield of AI that mimics the way humans learn by using data and algorithms to refine their performance over time. However, deep learning, a type of machine learning, consists primarily of neural networks with three or more layers. Machine learning and deep learning are both methods and tools that help us get closer to true artificial intelligence. The most important areas of use for artificial intelligence in the healthcare sector include diagnosis and treatment recommendations, patient engagement, and back-end processes. Despite the many instances in which AI can perform healthcare tasks at par with or above human levels.

Women between the ages of 12 and 50 are most vulnerable to developing polycystic ovary syndrome (PCOS). The cause of polycystic ovary syndrome (PCOS) remains unknown. Weight gain, irregular menstrual periods, excess body hair, hair loss or male-pattern baldness, acne or oily skin, and sporadic infertility are all common symptoms of polycystic ovary syndrome (PCOS). Formerly, the majority of experts believed that polycystic ovary syndrome was an endocrine disorder, but then studies now indicate it to be a metabolic, hormonal, and psychological illness, which has an impact on a patient's general well-being. Thus, early detection of this will assist the female population by enabling them to manage the mental stress that is frequently disregarded with polycystic ovary syndrome. We may utilise a variety of machine learning techniques to detect PCOS. This article reviewed several publications on PCOS detection from a wide range of journals.

II. LITERATURE REVIEW

To diagnose PCOS, we can use various machine learning approaches. In the article titled "Comparative Analysis of Machine Learning Algorithms in diagnosis of Polycystic Ovarian Syndrome," In their article, Malik Mubasher Hassan and Tabasum Mizra describe the steps they used to diagnose PCOS from patient clinical data using machine learning techniques such as Support Vector Machine, CART, Naive Bayes Classification, Random Forest, and Logistic Regression. Their paper's major goal is to use well-known machine learning algorithms on samples of random data to diagnose PCOS based on clinical symptoms associated with the condition. They then assess the effectiveness of several algorithms to choose the one that performs the best.

Coming to the paper "Detecting PCOS using Machine Learning" by Narmada Tanwani, a model is created based on the causes and symptoms to accept them as features and to output the presence or absence of this condition. In this research work a few machine learning models are created to ascertain whether PCOS is present. K-Nearest Neighbor (K-NN) and Logistic Regression are two supervised machine learning techniques that are employed since the dataset does classify whether the condition is present or not.

In their paper titled "PCOcare: PCOS Diagnosis and Prediction Using Machine Learning Algorithms," Vaidehi Thakre and Shreyas Vedpathak offer a method that can aid in the early diagnosis and prognosis of PCOS therapy utilizing an optimum and minimal set of parameters. Five distinct machine learning classifiers (Random Forest, SVM, Logistic Regression, Gaussian Naive Bayes, and K Neighbours) have been used to determine whether or not a woman has polycystic ovary syndrome.

©IJRASET: All Rights are Reserved

III. METHODOLOGY

A reliable methodology is the backbone of any fruitful investigation. The methodology of a research paper details the steps taken to conduct the research. The reader can use these data to evaluate the strategy's credibility and precision. Data collection, data pre-processing, and classification are the three steps adopted by *Malik Mubasher Hassan and Tabasum Mizra's* used for their study, "*Comparative Analysis of Machine Learning Algorithms in diagnosis of polycystic ovarian syndrome*." To diagnose PCOS, there were 42 independent variables used to train five machine learning algorithms on random data samples. There are two alternative outcomes for the dependent variable PCOS, which is significantly inversely related to these independent factors.: "Yes" or "No." While validating models, performance assessment criteria like as accuracy, precision, F-statistics, and recall are employed.

While in the paper "Detecting PCOS using Machine Learning" she had gone through the methodology, which includes the steps *Data collection*- samples from various hospitals across Kerala, India, *Data Analyzing*- understanding the collected data and their attributes, *Feature selection*- selecting the appropriate attribute as a feature. The weights of the characteristics are discovered using the filter approach in order to ascertain which of them have a strong link with the aim., *Fitting into the model*-K-NN and Logistic Regression are the two models utilised for supervised machine learning. The accuracy of K-NN depends upon two factors: The value of 'K' and the number of selected features. In logistic Regression as well, the accuracy depends on the number of features selected, *Making Predictions*- Predictions are made with a testing set using the provided models Evaluation- model is evaluated using the criteria Precision, Recall, and Support

As per Vaidehi Thakre and Shreyas Vedpathak in their study "PCOcare: PCOS Diagnosis and Prediction Using Machine Learning Algorithms," no data preprocessing was necessary because the dataset was already cleansed. Statistical aid has been used to identify key traits (Chi-Square Method). From a total of 41 features in the dataset, only the top 30 were selected using the CHI Square method. In statistics, the chi-square test is used to evaluate the degree of dependence between two events. A feature that is more reliant on the response and has a higher chi-square value may be chosen for model training. Following feature selection, they put into practise several machine learning techniques, such as the Random Forest Classifier, which integrates the output from numerous decision trees to arrive at a conclusion. Support Vector Classifiers are tools for outlier identification, regression, and classification issues. Gaussian Naive Bayes uses the Bayes theorem to determine which class the provided data would belong in and to compute the probability. After putting the machine learning algorithms into practise, the performance metrics—precision, recall, and F score—for each model was evaluated based on the test data.

IV. RESULT

Through Malik Mubasher Hassan and Tabasum Mizra's analysis suggested that the Random Forest algorithm performed best in the PCOS diagnosis according to the performance validation metrics recall, accuracy, precision, and F- statistics, with an accuracy of 96%, followed by SVM with a good accuracy. As a result, it is determined that the Random Forest algorithm is the most appropriate method for PCOS diagnosis using the provided data. The use of various or sizable data sets for illness diagnosis may be part of the study's future scope.

In the paper "Detecting PCOS using Machine Learning" by Narmada Tanwani, the two separate classifiers—linear and nonlinear—were compared to one another. KNN is a linear classifier, whereas the non-linear classifier is a model of logistic regression. To choose between the two models, the F1 score is useful. The F1 score for the Logistic Regression model is 0.92, whereas the F1 score for the KNN model is 0.90. As a result, the Logistic Regression model is chosen to identify if PCOS is present or absent.

According to the paper "PCOcare: PCOS Diagnosis and Prediction Using Machine Learning Algorithms," with an accuracy rate of 90.9%, the Random Forest Classifier was determined to be the most reliable and accurate of all. The suggested strategy can be used by both patients and doctors.

V. CONCLUSION

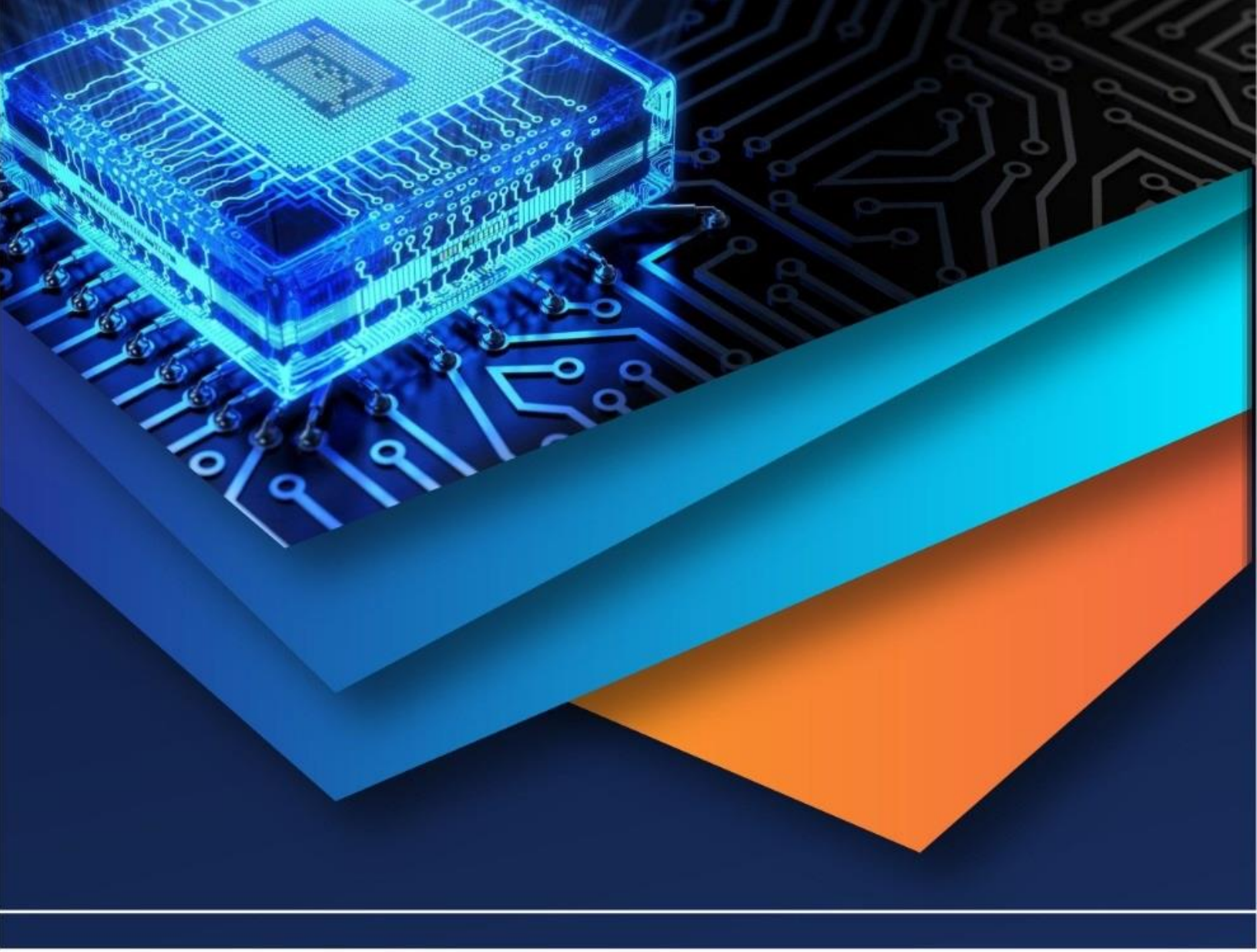
Polycystic ovarian syndrome is a complex endocrine disorder that affects females during their reproductive years. The Random Forest classifier is the best algorithm to predict this condition, according to the examination of the aforementioned articles. It affects the majority of women in society, and if left untreated or given the wrong diagnosis, these women are more likely to develop cardiovascular disease, type 2 diabetes, ovarian and uterine cancer, as well as reproductive issues. A doctor can screen new patients with basic information and give priority to treating persons with PCOS before seeing patients who do not have the ailment; therefore the proposed approach can be used for the early diagnosis of this sickness and will be helpful for both patients and doctors as well.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

REFERENCES

[1] Malik Mubasher Hassan and Tabasum Mirza. Comparative Analysis of Machine Learning Algorithms in Diagnosis of Polycystic Ovarian Syndrome. International Journal of Computer Applications 175(17):42-53, September 2020. DOI:10.5120/ijca2020920688
[2] Elmannai, Hela, Nora El-Rashidy, Ibrahim Mashal, Manal Abdullah Alohali, Sara Farag, Shaker El-Sappagh, and Hager Saleh. 2023. "Polycystic Ovary Syndrome Detection Machine Learning Model Based on Optimized Feature Selection and Explainable Artificial Intelligence" *Diagnostics* 13, no. 8: 1506. https://doi.org/10.3390/diagnostics13081506
[3] Tanwani, Namrata. (2020). Detecting PCOS using Machine Learning. 10.13140/RG.2.2.10265.24169.
[4] Thakre, V., Vedpathak, S., Thakre, K. & Sonawani, S. S. (2020). PCOcare: PCOS Detection and Prediction using Machine Learning Algorithms. Bioscience Biotechnology Research Communications, 13:240–244. doi: 10.21786/bbrc/13.14/56
[5] Vineetha Sankar, P., Sreekumar, K. (2022). Utilizing the Data Mining Techniques for Obesity Prognosis Based on Eating and Lifestyle Routines of Adolescents and Adults. In: Bianchini, M., Piuri, V., Das, S., Shaw, R.N. (eds) Advanced Computing and Intelligent Technologies. Lecture Notes in Networks and Systems, vol 218. Springer, Singapore. https://doi.org/10.1007/978-981-16-2164-2_30
[6] G. N. Allahbadia and R. Merchant, "Polycystic ovary syndrome and impact on health," Middle East Fertil. Soc. J., vol. 16, no. 1, pp. 19–37, 2011, doi: 10.1016/j.mefs.2010.10.002
[7] Escobar-Morreale H.F. Polycystic ovary syndrome: Definition, aetiology, diagnosis and treatment. Nat. Rev. Endocrinol. 2018; 14:270–284. doi: 10.1038/nrendo.2018.24.
[8] G. P. Rédei, "Polycystic Ovarian Disease (SteinLeventhal syndrome)," in Encyclopedia of Genetics, Genomics, Proteomics and Informatics, Springer Netherlands, 2008, pp. 1528–1528.
[9] Tiwari S., Kane L., Koundal D., Jain A., Alhudhaif A., Polat K., Zaguia A., Alenezi F., Althubiti S.A. SPOSDS: A smart Polycystic Ovary Syndrome diagnostic system using machine learning. Expert Syst. Appl. 2022; 203:117592. doi: 10.1016/j.eswa.2022.117592.
[10] G. N. Allahbadia and R. Merchant, "Polycystic ovary syndrome and impact on health," Middle East Fertil. Soc. J., vol. 16, no. 1, pp. 19–37, 2011, doi: 10.1016/j.mefs.2010.10.002.
[11] S. M. Sirmans and K. A. Pate, "Epidemiology, diagnosis, and management of polycystic ovary syndrome," Clin. Epidemiol., vol. 6, no. 1, pp. 1–13, 2013, doi: 10.2147/clep.s37559.
[12] Barber, T. M., & Franks, S. (2021). Obesity and polycystic ovary syndrome. Clinical endocrinology, 95(4), 531-541.
[13] Bharati, S., Podder, P., & Mondal, M. R. H. (2020, June). Diagnosis of polycystic ovary syndrome using machine learning algorithms. In 2020 IEEE Region 10 Symposium (TENSYMP) (pp. 1486-1489). IEEE.
[14] Chauhan, P., Patil, P., Rane, N., Raundale, P., & Kanakia, H. (2021, June). Comparative analysis of machine learning algorithms for prediction of pcos. In 2021 International Conference on Communication information and Computing Technology (ICCICT) (pp. 1-7). IEEE.
[15] Watson, S. (2019). Polycystic Ovary Syndrome (PCOS): Symptoms, Causes, and Treatment.
[16] 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.
[17] Mehreen, T. S., Ranjani, H., Kamalesh, R., Ram, U., Anjana, R. M., & Mohan, V. (2021). Prevalence of polycystic ovarian syndrome among adolescents and young women in India. Journal of Diabetology, 12(3), 319-325.
[18] Bhardwaj, P., & Tiwari, P. (2022). Manoeuvre of Machine Learning Algorithms in Healthcare Sector with Application to Polycystic Ovarian Syndrome Diagnosis. In Proceedings of Academia-Industry Consortium for Data Science: AICDS 2020 (pp. 71-84). Singapore: Springer Nature Singapore.
[19] Watson, S. (2019). Polycystic Ovary Syndrome (PCOS): Symptoms, Causes, and Treatment.
[20] Neuzil, A. (2014). What is polycycstic ovary syndrome (pcos). URL: https://www. austinfitmagazine. com/November-2014/What-is-Polycycstic-Ovary-Syndrome-PCOS.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)