International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 270-275
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0037

A Comprehensive Machine Learning-Based Predictive Platform for Early Detection and Analysis of Infertility in Both Male and Female Using Ensemble Learning Techniques

Jeevitha C S[1], Sampada S Kapse[2], Ankita[3], Firdos Begum[4] , Lakshmi Durga[5]
[1]UG Scholar, Dept. of ISE, AMC Engineering College., Bangalore, Karnataka, India
[2]Assistant Professor, Dept. of ISE, AMC Engineering College, Bangalore, Karnataka, India
[3]UG Scholar, Dept. of ISE, AMC Engineering College., Bangalore, Karnataka, India
[4]UG Scholar, Dept. of ISE, AMC Engineering College., Bangalore, Karnataka, India
[5]UG Scholar, Dept. of ISE, AMC Engineering College, Bangalore, Karnataka, India
Emails: *jeevithacsgowda@gmail.com[1], Sampada.kapse@amceducation.in[2], ankitareddybogale@gmail.com[3], firdosbegum07865@gmail.com[4], gbhavani2164@gmail.com[5]*

Abstract
The inability to conceive after 12 months of consistent, unprotected sexual activity is known as infertility, a common reproductive health problem. Millions of couples worldwide are impacted by it, and it can be caused by issues pertaining to the male, female, or both partners. Genetic disorders, hormonal imbalances, structural abnormalities, infections, and lifestyle factors like poor diet, smoking, alcohol consumption, and stress can all contribute to male infertility. A thorough medical history, physical examination, semen analysis, and additional diagnostic testing as needed are typical components of evaluation. Ovulatory conditions like polycystic ovary syndrome (PCOS), fallopian tube damage, uterine abnormalities, endometriosis, or age-related decline in fertility are frequently linked to female infertility. Hormonal evaluations, ultrasound imaging, and specialized procedures like laparoscopy in certain situations are among the diagnostic techniques used for women. Depending on the underlying cause, treatment options for infertility may involve medication, surgery, lifestyle changes, or assisted reproductive technologies like intracytoplasmic sperm injection (ICSI) and in vitro fertilization (IVF). Accurate diagnosis and successful treatment depend on a comprehensive assessment of both partners. Reproductive medical advancements keep improving treatment results, giving infertile couples more hope and a higher standard of living.

Keywords: *Assisted Reproductive Technology (ART); Diagnosis and Treatment; Female Infertility; In-vitro Fertilization; Reproductive Health.*

1. Introduction
The inability to conceive after 12 months of consistent, a thorough medical history, physical examination, semen analysis, and additional diagnostic testing as needed are all common components of evaluation. Ovulatory conditions like polycystic ovary syndrome (PCOS), fallopian tube damage, uterine abnormalities, endometriosis, or age- related decline in fertility are frequently linked to female infertility. Hormonal evaluations, ultrasound imaging, and specialized procedures like laparoscopy in certain situations are among the diagnostic techniques used for women. unprotected sexual activity is known as infertility, a common reproductive health problem. Millions of couples worldwide are impacted by it, and it can be caused by issues pertaining to the male, female, or both partners. Genetic disorders, hormonal imbalances, structural abnormalities, infections, and lifestyle factors like poor diet, smoking, alcohol consumption, and stress can all contribute to male infertility.

2. Methodology
The first step in this study is gathering clinically

significant male and female fertility data from trustworthy medical databases and medical records. To handle missing values and minimize noise, the data is meticulously cleaned, normalized, and preprocessed. Statistical analysis and feature selection techniques are used to identify significant factors influencing infertility (Figure 1). To identify various patterns in the data, several machine learning models are trained separately.

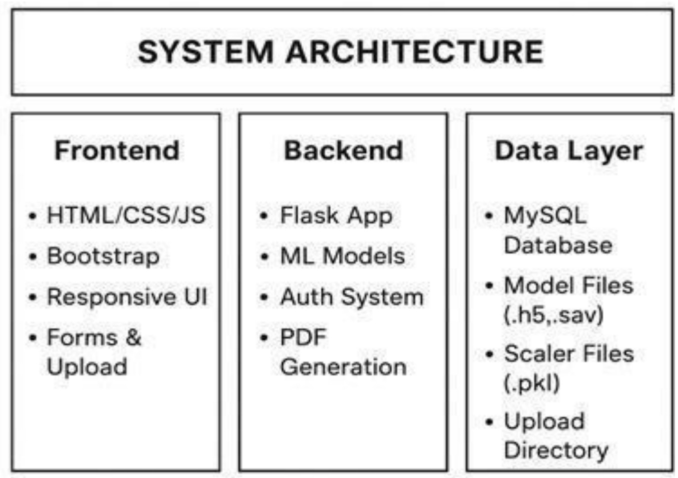

Figure 1 System Architecture

2.1. Frontend (User-Interface Layer)
The platform's front-end is created with HTML, CSS, and JavaScript to offer a well-organized layout, eye- catching styling, and seamless user interaction. In order to guarantee a contemporary, responsive design that easily adjusts to various devices and screen sizes, Bootstrap is integrated.

2.2. Backend (Logic Layer)
Flask, a lightweight and adaptable Python web framework that effectively handles server-side operations, is used to implement the system's back-end. It manages user requests for prediction and analysis and hosts the trained machine learning models. To regulate user logic and platform access, a secure authentication system is integrated. Furthermore, users can download comprehensive reports and results in an easy-to-read format thanks to the system's support for PDF generation.

2.3. Data Layer (Storage and Model Files)
To safely store structured data, including user information and prediction outcomes, the system makes use of a MySQL database. To facilitate effective loading and reuse during analysis, trained machine learning model files, such as those in the.h5 and .sav formats, are kept apart. Consistent data preprocessing throughout the training and prediction phases is ensured by scaler files stored in the.pkl format. Additionally, user-uploaded files or datasets needed for processing are securely stored in a dedicated upload directory (Figure 2).

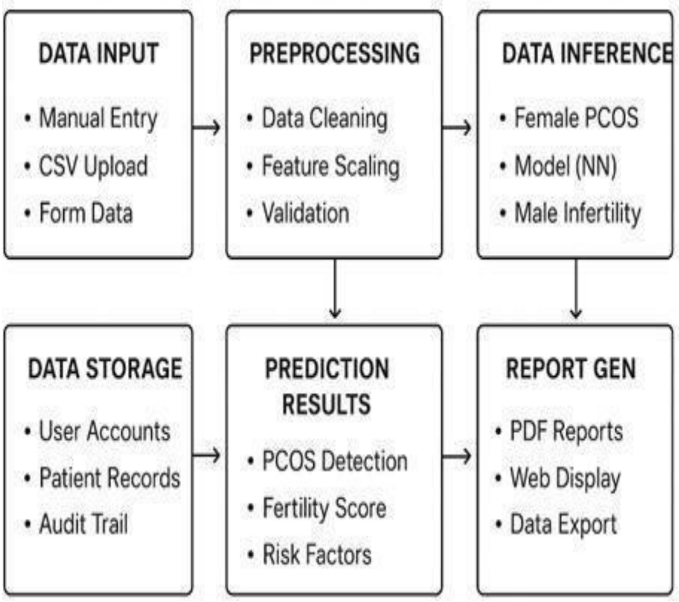

Figure 2 Methodology

2.4. Data Input
To provide users with flexibility and ease of use, the data input module collects data from various sources. Through user-friendly forms, it permits manual data entry, allowing people to enter information straight into the system. For bulk data processing, users can also upload datasets in CSV format. Data collection is further streamlined by web-based form submissions, which make the input process as a whole easy, accessible, and effective.

2.5. Preprocessing
By enhancing its quality and dependability, the preprocessing step gets the gathered data ready for efficient modeling. It entails cleaning the data to get

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 270-275
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0037

rid of mistakes, inconsistencies, and missing values that might affect how accurately predictions are made.

2.6. Data Inference

The trained machine learning models are used in the data inference stage to evaluate the processed input data and produce insightful predictions. By identifying pertinent health patterns and indicators, neural network-based models are used to predict female PCOS. In order to assess the factors that contribute to reproductive health, an analysis of male infertility is conducted concurrently.

2.7. Data Storage

To guarantee dependability and privacy, the data storage components safely preserve all system records. It provides safe and controlled access by storing user accounts and profile data in an orderly fashion. In order to facilitate precise analysis and future reference, patient medical records are maintained in an appropriate manner. Additionally, audit trails are kept to monitor data access and changes, improving system security, accountability, and transparency.

2.8. Prediction Results

Clear and significant results produced by the system are displayed in the prediction results module. In order to highlight possible issues with female reproductive health, it shows the PCOS detection status. In addition, a risk index or fertility score is offered to assist users in comprehending the overall degree of infertility risk.

2.9. Report Generation

Prediction results are transformed into well-structured and comprehensible outputs by the report generation module. It generates PDF reports that can be downloaded and accessed offline for sharing or future reference. Additionally, results are shown directly on the web interface for quick interpretation and real-time viewing. Furthermore, the system permits data export in appropriate formats, allowing for additional analysis or tool integration when needed.

3. Figures

Figure 3 Sign Up Page

The secure login screen for the infertility analysis system is shown in Figure 3. To access the platform and its features, users must input their username and password. The login process is simple and easy to use thanks to the design's organization and simplicity. For new users without an account, there is also a sign-up option.

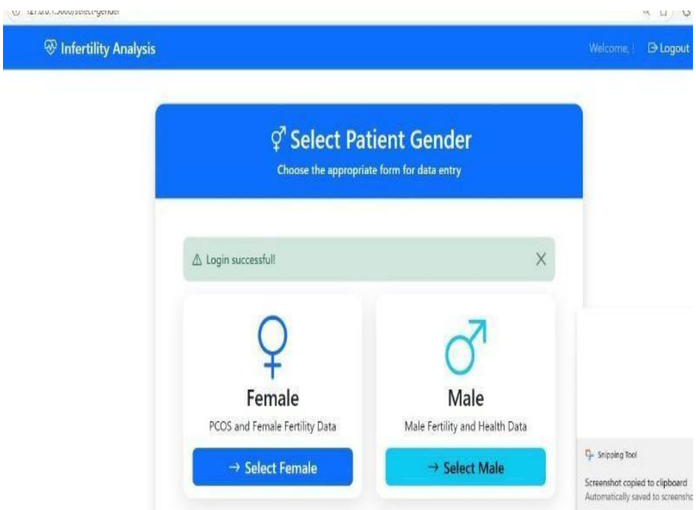

Figure 4 Gender Selection Page

IRJAEH

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 01 January 2026
Page No: 270-275
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0037

The gender selection interface of the infertility analysis system is displayed in Figure 4. Depending on the kind of fertility analysis they wish to conduct, users can select either male or female.

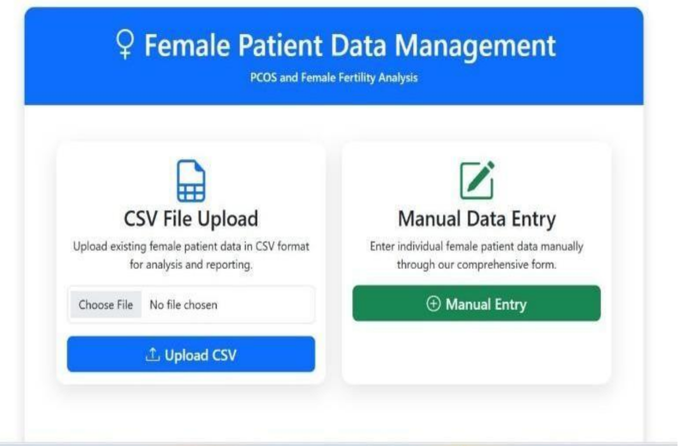

Figure 5 Patient Data Entry Page

The infertility analysis system's female patient data entry section is displayed in Figure 5. It enables users to enter patient data manually via a form or by uploading a CSV file. Figure 6 shows the report generated.

PCOS PATIENT DATA ANALYSIS REPORT

Generated by: js | Date: 2025-11-05 14:54:08

Attribute	Record 1
Age (yrs)	26
Weight (Kg)	49.792563499
Height(Cm)	152.1764350841
BMI	21.8971627347
Blood Group	11
Pulse rate(bpm)	72
RR (breaths/min)	20
Hb(g/dl)	11
Cycle(R/I)	4
Cycle length(days)	5
Marraige Status (Yrs)	2
Pregnant(Y/N)	0
No. of abortions	0
I beta-HCG(mIU/mL)	1.99
II beta-HCG(mIU/mL)	1.99
FSH(mIU/mL)	5.7
LH(mIU/mL)	4.5
FSH/LH	1.27
Hip(inch)	32
Waist(inch)	28

Figure 6 Report Generation

4. Results and Discussion

4.1. Results

Following appropriate preprocessing and feature scaling, the results of the suggested infertility prediction system demonstrate that all three machine learning models operated consistently and effectively. The neural network-based PCOS prediction model for females had the highest accuracy of roughly 89.2%, successfully differentiating between PCOS and non-PCOS cases with high sensitivity and specificity. It produced fast predictions within the Flask-based application and was particularly impacted by important hormonal characteristics like the follicle count and LH/FSH ratio. Using a Random Forest ensemble, the male lifestyle-based fertility model demonstrated balanced precision and recall across fertility classes, with an accuracy of roughly 85.7. Prediction was significantly influenced by lifestyle factors such as age, stress levels, alcohol consumption, smoking, and physical activity. The ensemble approach prevented overfitting while preserving quick inference. Sperm motility emerged as the most significant parameter in the semen analysis model, which was implemented using a Decision Tree classifier and recorded an accuracy ranging from 82% to 86%. It performed especially well for normal fertility cases. The model produced extremely accurate and quick predictions, though it was marginally less successful in cases that were slightly altered. Overall, the comparative analysis shows that while ensemble and tree-based approaches worked well for structured lifestyle and semen data, deep learning was best for complex hormonal patterns in PCOS detection. The system is appropriate for early infertility screening and supportive clinical decision-making because of its strong predictive ability and practical efficiency.

4.2. Discussion

According to the results discussion, the suggested infertility prediction system successfully uses machine learning techniques to support both male and female fertility assessments. With accuracy levels between 80% and 90%, which are thought to be appropriate for medical decision-support applications, all three models showed dependable performance and remained stable under various testing circumstances. The system's real- time usability is one of its main advantages; even when processing large inputs like CSV files, predictions were produced in less than a second via the Flask-based interface. From a clinical standpoint, each module has significant value: the male lifestyle-based model offers useful information that can promote healthier fertility- related behaviors, while the PCOS prediction model helps identify women who might benefit from testing. The model provides an early warning of possible reproductive issues. Additionally, the system is built with scalability in mind, facilitating future cloud deployment and making it simple to integrate additional prediction modules. The findings do, however, also point out some drawbacks, such as the requirement for more extensive and superior medical datasets in order to improve accuracy and dependability. Crucially, even though the system offers useful predictive insights, it must be used in conjunction with professional medical interpretation and is not meant to take the place of clinical diagnosis. All things considered, the system has great potential to assist clinics, telemedicine services, fertility counseling, and personalized health monitoring by providing quick, data-driven predictions. This will help to minimize diagnostic delays and encourage early detection of fertility- related problems.

Conclusion

The discussion demonstrates how well the infertility prediction system supports machine learning-based fertility assessments for both men and women. With accuracy ranging from 80% to 90%, all three models produced good results that are suitable for medical support systems. Even when several records are uploaded at once, the system is simple to use and yields results quickly—typically in less than a second. Each model has practical significance: the semen analysis model aids in the early detection of reproductive issues, the male lifestyle model offers helpful recommendations for improving fertility through better habits, and the PCOS model helps identify women who might require additional hormonal testing. Because of its flexibility, the system can be expanded in the future by deploying it on the cloud and adding new prediction models. However, by utilizing more extensive and superior medical data, the accuracy can be increased. Additionally, the system is designed to assist physicians rather than to take the place of medical diagnosis. Overall, by offering quick and early insights into fertility-related issues, this system can be helpful for clinics, telemedicine platforms, and fertility counseling.

Acknowledgement

The author would like to thank their institution, their faculty advisors, and the technical team for their support throughout this project They also want to recognize the creators of the ICBHI dataset for supplying the important resources that were crucial for their research

References

[1]. A. O. Akinola, M. A. Ibrahim and H. A. Abdulrahman, “Machine Learning-Based Diagnosis of Polycystic Ovary Syndrome Using Clinical Data,” IEEE Access, vol. 8, pp. 18953–18965, 2020.

[2]. G. S. Sundar, R. P. Vidyadharan and S. Nair, “Artificial Intelligence Models for Predicting Female Fertility Disorders,” in Proc. IEEE Int. Conf. Bioinformatics and Biomedicine (BIBM), pp. 112–118, 2021.

[3]. S. R. Jadhav and P. M. Patil, “Deep Learning Approach for Early Detection of Polycystic Ovary Syndrome Using Hormonal Patterns,” IEEE Trans. Artificial Intelligence, vol. 2, no. 4, pp. 301–309, 2021.

[4]. A. M. Fernandes, V. R. Carvalho and T. C. Silva, “Predicting Male Fertility Using Ensemble Machine Learning Techniques,” in Proc. IEEE Int. Conf. Healthcare Informatics (ICHI), pp. 145–152, 2020.

[5]. F. A. Shaikh and M. A. Pathan, “Decision Tree- Based Semen Quality Classification for Infertility Prediction,” IEEE Int. Conf. Computing Methodologies (ICCM), pp. 215–220, 2021.

[6]. M. El-Gayar, A. Timsina and S. Nawar, “A

Data Mining Framework for Predicting Fertility Outcomes Using Clinical Parameters," IEEE Access, vol. 7, pp. 14067–14076, 2019.

[7]. R. K. Mishra, S. Agrawal and R. Tripathi, "A Hybrid Machine Learning System for Reproductive Health Risk Prediction," in Proc. IEEE Int. Conf. Emerging Technologies in Computer Engineering, pp. 227–234, 2020.

[8]. N. Gupta and S. K. Singh, "AI-Based Clinical Decision Support System for Female Reproductive Health," IEEE Sensors Journal, vol. 21, no. 10, pp. 11545–11553, 2021.

[9]. P. Thakur, J. Bose and A. H. Mir, "A Predictive Analytics Model for Infertility Assessment Using Lifestyle Factors," in Proc. IEEE Int. Conf. Intelligent Systems Design and Applications (ISDA), pp. 442–448, 2020.

[10]. S. Khan and T. Hussain, "Application of Machine Learning in Healthcare for Early Disease Prediction: A Case Study on PCOS," IEEE Int. Conf. Machine Learning and Applications (ICMLA), pp. 985–990, 2019.

[11]. D. Patel, R. Mehta and J. Soni, "Medical Data Classification Using Machine Learning for Fertility Analysis," IEEE Int. Conf. Data Engineering and Communication Technology, pp. 410–416, 2021.

[12]. M. H. Yousuf, A. Farooq and K. Abbas, "Predicting Hormonal Imbalances Using AI Models for Reproductive Health," IEEE Access, vol. 9, pp. 126701–126712, 2021.