Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 47 No. 03
March 2026

Ensemble Machine Learning Methods for Hypertension Scaling in Pregnant during Trimester Periods

Bh. V. RamaKrishna[1], Shreeja Vydyam[2], Eruvanti Madhumitha[3]

[1] Associate Professor, Vignan Institute of Technology and Science, Deshmukhi, Telangana, India

[2,3] Dept. of AI & DS, Vignan Institute of Technology and Science, Deshmukhi, Telangana, India

Abstract

In recent years more interest of research is taking place for analyzing risks for women during pregnancy period. Increased machine learning techniques application over medical data to identify risk factors during trimester periods provide an insight to knowledge patterns supporting good decision making. In this paper we present ensemble methods for hypertension scaling and classification of clinical features. The adoption of AI techniques and transformer encoding methods are under research. Hypertension severity estimation and trimester based mental health data is the focus carried over in this research to generate decision summaries for pregnant health care guidance in perspective of both mother and child growth harmony. Various models are adopted for improved interpretability analysis. The application of transformer learning mechanisms encouraged in this research to introduce hybrid ensemble models for hyper scaling pregnancy trimester risks.

Keywords: Machine Learning, Hypertension, Risk Analysis, Trimesters

1. Introduction

In developing country like India pregnant women living in rural areas are 9 - 30% more suffering from hypertensive disorders. The lower access to healthcare socioeconomic barriers increasing the morbidity and mortality rates in rural communities [4]. During second and third trimesters various health factors including blood pressure increases the chance of preeclampsia, gestational hypertension and low birth weights, the standard threshold for hypertension during pregnancy 140/90mmHg, also 130/80mmHg identified as risk according to fetal assessments [6]. Electronic Medical Records support global data analysis at quick times to reduce maternity risks and gynec problems [2]. Some ensemble methods enhance the accuracy of risk prediction and data analysis quality to 92% nowadays [10]. Many researchers focused over integration of AI and ML technologies to improve the health monitoring and risk prediction assessment [8]. Modern medical healthcare environments are highly sensor equipment integrated. AI promises to fill the gap between data monitoring and assessment techniques with personalized recommendations [9].The wearable devices are in demand for real-time diagnosis of women health metrics with precision [14].

2. Hyperpiesia In Pregnancy

In general pregnancy risks are higher in rural women due to limited medical service access and lack of efficient diagnostic centers. They face higher maternal morbidity compared to urban women. There is increased chance of risks like preterm birth and infant mortality. The rural women are surrounded with socioeconomic and insurance barriers to support expensive medical checkups. The prenatal care is delayed for them causing early high risk detection failures.Some common risks, which are specified in Table 1, must be carefully tracked and mitigated for each trimester to ensure timely intervention.

Table1. Trimester Hypertension Risks

Trimester Period	HypertensionRisks
First	Preeclampsia risk Poor Placement risk Fetal growth restriction Early miscarriage risk Placentaldevelopment risk
Second	Gestational hypertension risk Miscarriage risk Preeclampsia risk Fetal growth restriction Preterm birth risk
Third	Severe Preeclampsia Eclampsia risk HELLP syndrome

	Placental abruption

Table2. Natural Diet for Pregnant

Food	Benefits
Cow Milk	Nourishes Fetal Health Bone Density Healthy Heart Vaginal Infections resistance
Ghee	Fetal Tissues growth Aids digestion Brain development Preterm birth risk reduction
Whole Grains	Protein support Fetal health Fetal immunity Brain growth
Fruits & Vegetables	Immunity Nutritional Vitamins Fetal Health Tissue building
Herbs	Resistance to Nausea, Migraine Reduce Vomiting Body Pains Aid Digestion Reduce bloating Iron support

In rural areas pregnant women rely on traditional and obtainable resources for health. In Telangana and Andhra-Pradesh '*Garbini Paricharya*' and '*Prenatal Yoga*' is a recommended framework by Gram Panchayath suitable accessibility rural structures. Ayurveda from generations support human prosperity by promoting vitality and immunity improvement with natural herbs and plant products without side effects.Table 2 explores some natural vegetarian diet and its role in building maternal health.

Table3. Hyperpiesia Symptom and Effects

Symptom	Effects
Dehydration	Severe fluid loss (oliguria) Orthostatic hypotension Intravenous fluid loss
Malnutrition	Skin Break(Ketonuria) Nutritional Deficiency
Electrolyte Imbalance	Vomiting Neuromuscular problems Cardiac Arrhythmias
Wernicke Encephalopathy	Brain growth problems Vision problems
Gastrointestinal Trauma	Reflux Vomiting Neuro disorder Abdominal problems

Many health issues often caused with hyper tension in women during trimesters. The Psycho-Social abilities are highly influenced with 'Hyperpiesia'. Table 3 shows the various symptoms rise due to hyper tension states and their effects over both mother and baby during trimester periods. Regular health checkups and balanced dietary keeps maternal body healthy. Some lifestyle modifications also needed for both fetus and mother health such as walking, prenatal yoga, swimming, meditation and pranayama. Avoid smoking and consuming alcohol.

3. Machine Learning Algorithms

We applied wide range of efficient machine learning techniques over training data related to pregnant women health records. These approaches are helpful to analyze the metrics accurately for projecting knowledge patterns which are used for predicting various risk factors during trimesters. Also provide recommendations for women safety during pregnancy time.

A. Association Rule Mining (ARM)

Association rule mining can be applied to pregnant women healthcare data to study the effect of hypertension during trimesters. For identifying frequent patterns and their relationships association mining helps a lot. In our research work various attribute combinations linked to hypertension and cardiovascular conditions in pregnants. The feature set considered included dietary habits, genetic disorders, hereditary and pain symptoms. ARM use Apriori algorithm to build dependent factors with estimating support and confidence rates.

K-Nearest Neighbors (KNN)

Applied in classification of medical image data, which is collection of continuous data. Based on data points distribution it classifies using a non-parametric approach that takes majority class of close neighbors. To place new data into related category by comparing it with current data. It uses Euclidean distance to classify.

Euclidean distance is given by

$$D = \left[\sum_{j=1}^{n}\left(u_j - v_j\right)^2\right]^{05} \qquad (1)$$

B. Support Vector Machine (SVM)

A supervised learning algorithm best suitable for classification of data using hyper planes. Simple to implement and most effective to handle large dimensional data spaces.

C. Decision Tree (DT)

Another supervised learning technique used for classification and regression tasks. Robust in handling multidimensional data. Supported with strong metrics to maintain accuracy and efficiency in classification. Can also be biased if dominating classes increases.

Entropy of Data is given by

$$E(Iins) = -\sum_{j=1}^{n} p_{a_j} \log_2 p_{a_j} \quad (2)$$

Information Gain is given by

$$Gain = E(Ins) - \sum_{i1}^{n} \frac{|Ins_i|}{|Ins|} E(Ins_i)(3)$$

D. Random Forest (RF)

An approach that creates many choice trees and classify the data according to Aggregates of predictions. It is robust to over fitting and provides high accuracy.

E. Gradient Boosting Machine (GBM)

A sequential boosting technique that optimizes the model by minimizing a loss function. It is effective for handling complex datasets. It is improving DTs by including poor learners gradually to form a strong prediction model. It can deal a variety of data types and give robust predictions with excellent performance.

F. XG-Boost

It is said to be an advanced form of gradient boosting that utilizes regularization and parallel processing for better performance and efficiency. A powerful supervised learning algorithm ensembles decision trees sequentially to correct previous generation tree data. Well known for its speed and memory efficiency. It is highly suitable for applications like classification, regression and ranking of data. Each algorithm undergoes cross validation and hyper parameter tuning to get optimal performance. Evaluation metrics provide the analysis of model performance. This Systematic approach helps in choosing of best performing algorithm for prediction.

4. Methodology

The data for this work collected from Kaggle® and few datasets are manually created based on the local MHRs collected from ANGANWADI, PHC and CHC. The data collected in the surrounding villages of Hyderabad located at Telangana state of India. ARMOData.csv, MHHRPDataset.csv and MHRADataset.csv are the training data sets used in this work for various ensemble machine learning methods. Each data set approximately a collection of records ranged up to 1000 and features up to 10 – 15. Methodology comprises of many stages. We began by collecting the dataset. Then analyze the dataset using different metrics and algorithms. Exploratory Data Analysis (EDA) is done to analyze the data. Unknown values and Outliers are identified and fixed in the data preprocessing. Feature Engineering is done. We used robust scaler and Standard scaler for the transformation of data. Apply different algorithms on the training and test data.

Finally compare different algorithms for performance Evaluation. Use oversampling technology to identify effectiveness of values missing and imbalance of class.

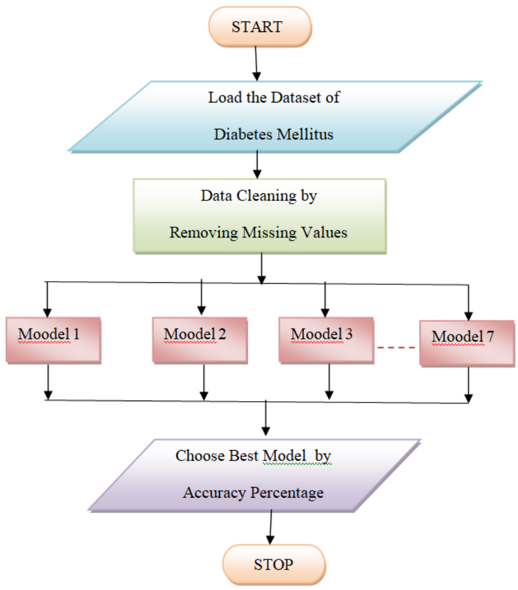

Fig.1 Flowchart of choosing Best Model.

The methodology demonstrates the machine learning algorithm's robust training on the clear and concise training data, to predict the unknown values. The aim is to harness the power of machine learning techniques for the classification of diseases in the healthcare sector. This methodology ensures the development of reliable predictive system for maternal hyper tension reduction.

5. Results And Discussions

Results are obtained are used to assess the findings. The outcomes serve as the measurements for accessing the algorithms' success. The average of the outcomes from k experiments, where k is the random selection of experimental data for k-fold cross validation, represents the model's performance.

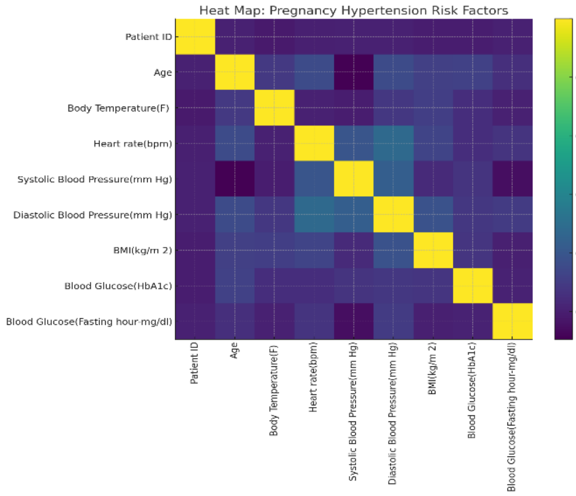

Fig.2. Heat Map for Hypertension Risk Analysis

The Heat map visualizes the knowledge patterns into warm and cool color zones. The features maintaining high intensity named as hot spots shows the severity of hypertension with relative cause factors.

Decision trees are useful to classify the hypertension training data to generate good classifiers which can classify future data with good precision and recall. The Bagged Decision Trees are modern solution to improve fastness in computation as well as preventing over fitting. Figure 3,4 are evidence for classification of vitamin deficiency risks impact over hypertension in maternal.

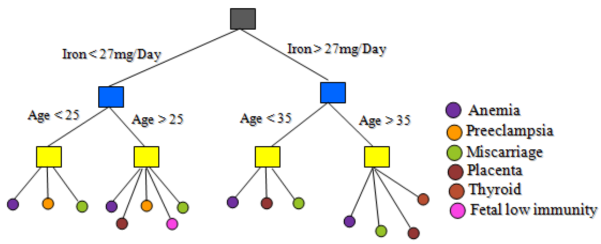

Fig.3. Decision Tree Iron deficiency risks

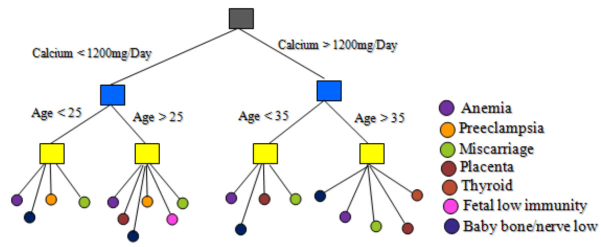

Fig.4. Decision Tree Calcium deficiency risks

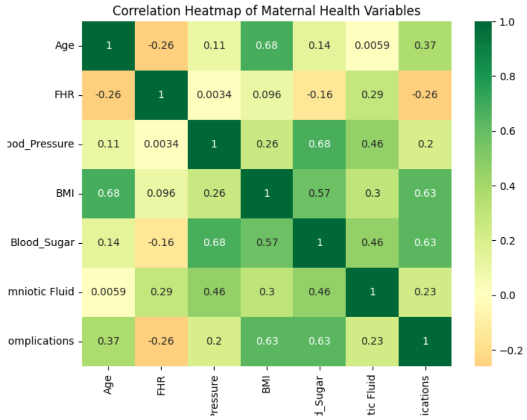

Fig.5. Correlation Analysis on MHV

The correlation matrix is a best statistical approach to determine the relationships and correlation factor rates among various features in data sets. Some major factors like Blood Pressure, BMI and Amniotic Fluid are highly influencing Fetal Growth and PHR from Figure 5. Another important visualization is Box-Plot Classification Report of Decision Tree Algorithm Variation machine learning models implemented. The results highlight the importance of the preprocessing techniques that are applied.

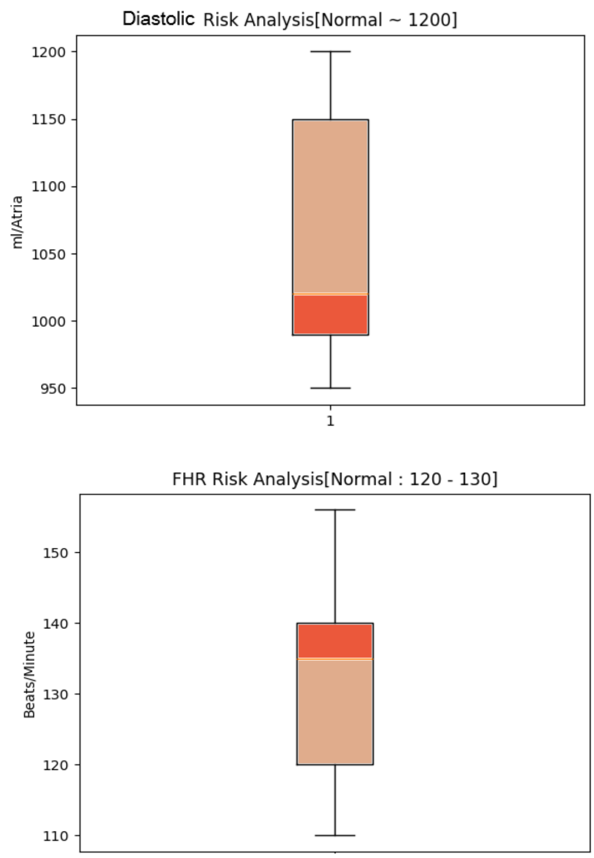

Fig.6. Risk Levels in PregnantsBox-Plots

Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 47 No. 03
March 2026

Box plots visually depict distribution of numerical data widely accepted to compare distributions between multiple groups. It shows the distribution of items in groups as quartiles. In Fig. 6 how FHR and Diastolic distributed over dimensional space and brisk represents the risk zone.

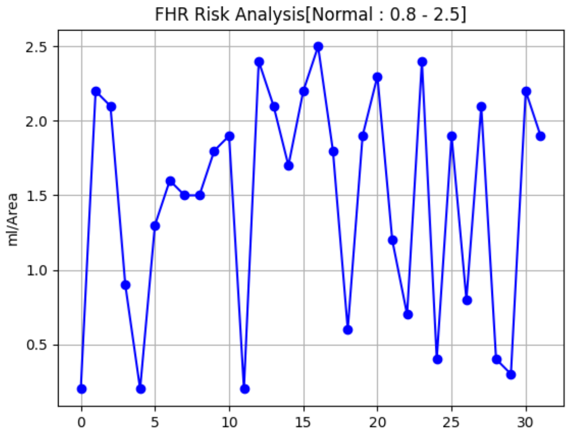

Fig.7. Risk Levels in PregnantsLine graph

The Association Rule Mining discovers relationships among features andfind associations based on support and confidence measures. The 'Lift' decides how much likely two features influence each other. Figure 8 represents the

A priori 1

Parameters

A-Priori parameters	
Support min	0.43
Confidence min	0.75
Max rule length	4
Lift filtering	1.10

Results

ITEMS

Transactions	129
Counting items	
All items	8
Filtered items	6
Counting itemsets	
card(itemset) = 2	5
card(itemset) = 3	1
Rules	
Number of rules	11

RULES

Number of rules : 11

N°	Antecedent	Consequent	Lift	Support (%)	Confidence (%)
1	"Preeclampsia=true" - "Diabetes=true"	"Miscarriage=true"	2.15000	46.512	100.000
2	"Preeclampsia=true" - "Miscarriage=true"	"Diabetes=true"	2.15000	46.512	100.000
3	"Diabetes=true"	"Miscarriage=true"	2.15000	46.512	100.000
4	"Miscarriage=true"	"Diabetes=true"	2.15000	46.512	100.000
5	"Diabetes=true"	"Preeclampsia=true" - "Miscarriage=true"	2.15000	46.512	100.000
6	"Miscarriage=true"	"Preeclampsia=true" - "Diabetes=true"	2.15000	46.512	100.000
7	"Diabetes=true"	"Preeclampsia=true"	1.20561	46.512	100.000
8	"Miscarriage=true"	"Preeclampsia=true"	1.20561	46.512	100.000
9	"Miscarriage=true" - "Diabetes=true"	"Preeclampsia=true"	1.20561	46.512	100.000
10	"Thyroid=true"	"Preeclampsia=true"	1.20561	61.240	100.000
11	"Obesity=true"	"Preeclampsia=true"	1.20561	51.938	100.000

Computation time : 16 ms.
Created at 20-11-2025 11.18.53 PM

Fig.8. Association Rule Mining for Pregnancy risks based on MHR

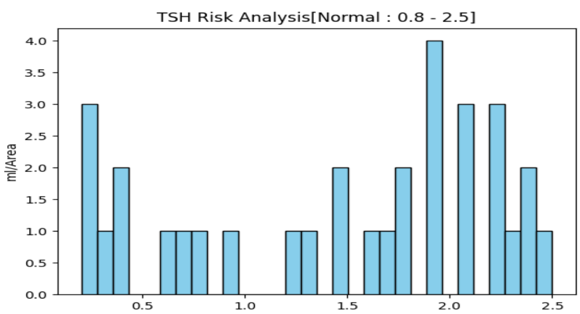

Fig.9. Histogram for Metarnal TSH statistics

When it comes to examine the frequency ranges among a large distribution one of the finest visualization handy tous is 'Histogram'. In our analysis we applied over TSH(Thyroid Stimulating Hormone) among pregnant women MHR statistics. The intention is to identify the patterns as shown in figure 9, for grouping rural women according to TSH release. Once the grouping done we further investigate the diet habits of groups and regular activities to give recommendations for women suffering from Low-TSH, High-TSH.

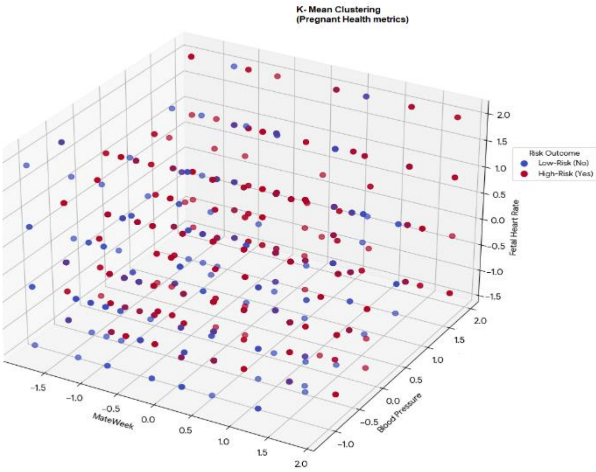

Fig.10. K-Mean Clustering of maternal health statistics.

Figure 10 pregnant health statistics based clustering with K-mean algorithm. A simple but effective clustering mechanism to handle multi-dimensional data in large volumes also. K-mean relies on 'Euclidean' distance measure and assumes clusters as equal sized. Before choosing K-mean we need to preprocess data using dimensionality reduction to get smooth and effective cluster boundaries. While applying over high dimensional data mini-batch approach improves efficiency of algorithm. Moreover K-means helps to discover hidden patterns by grouping data into smaller entities supporting better decision making in diversed applications.

Figure 10 shows the analysis conducted over pregnant risk metrics by considering features like *Mateweek*, *Blood Pressure*, *Fetal Heart Rate* to identify the risk severty groups. Here the groups are formed based on Low/High risk based with selected feature influences.

Table3. Hyperpiesia Symptom and Effects

Classification Method	Accuracy
Logistic Regression	75.0%
Decision Tree	89.5%
Random Forest	88.5%
Support Vector Machine	69.0%
K-Nearest Neighbors (KNN)	76.5%
Gradient Boosting	85.0%
XG – Boosting	83.6%
Naive Bayes	73.2%

By the result analysis over ensemble learning algorithms it is observed clearly for pregnant trimester data sets classification Decision Trees showing high accuracy up to 89.5% where Random Forests are helpful to large dataset classification with concurrent tree building nets. With an accuracy of 85%, the Gradient Boosting algorithm stands as the second-best choice and is highly suitable for pattern analysis due to its ability to capture complex relationships.The least accuracy with 69% SVM is computationally unadoptable for our datasets. Figure 11 depicts the overall hypertension scaling on rural maternal metrics which indicate that those working in environments with high intensive work suffering from hypertension compared to other. The Ethnic women are facing major societal restrictions and away from regular medical assistances depending on traditional medicine with unawareness of actual cause for risks are at top position in facing hypertension risks.

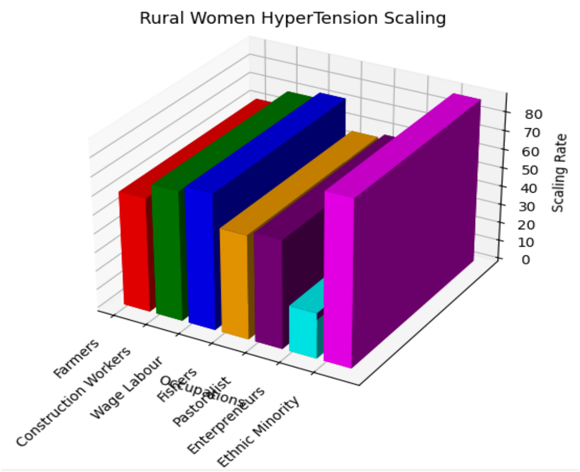

Fig.11. K-Mean Clustering of maternal health statistics.

6. Conclusion And Future Scope

In this research presented an ensemble machine learning–based framework for effective hypertension scaling in pregnant women across different trimester periods, addressing a critical challenge in maternal healthcare. By integrating multiple predictive models through ensemble strategies such as Random Forest, XG-Boost, and Ada-Boost, the proposed an approach for superior predictive accuracy, robustness, and generalization compared to individual base learners.The trimester based analysis enabled the model to capture the dynamic physiological variations in blood pressure patterns during pregnancy, allowing for more precise risk stratification of gestational hypertension and preeclampsia. Experimental results confirmed that ensemble learning significantly reduces misclassification errors, improves sensitivity for high-risk cases, and enhances overall clinical reliability, which are an essential requirement for early intervention and preventive maternal care. Furthermore, the proposed framework supports scalable deployment in real-world clinical decision support systems, especially in resource-constrained healthcare environments. By providing timely and accurate hypertension risk assessment, this work contributes to improved maternal andfetal outcomes and reinforces the role of intelligent data-driven systems in obstetric care. Future work may extend this study by incorporating longitudinal monitoring data, explainable AI techniques for clinical interpretability, and integration with wearable or IoT-based health monitoring systems to further strengthen personalized prenatal healthcare delivery.

References

[1] Katerina D. Tzimourta et. el., "Maternal Health Risk Detection: Advancing Midwifery with Artificial Intelligence", MDPI _ Healthcare Journal, Vol. 13, Issue - 7, ISSN: 2227-9032, PP. 1-21, 2025.

[2] Mohammad Mamun et. el., "Identification of Maternal Health Risk From Optimal Features Using Explainable Machine Learning", Wiley Publishers, Engineering Reports, Vol. 7, Issue -11, ISSN: 2577-8196,DOI:/doi.org/10.1002/eng2.70491, PP. 1-20, 2025.

[3] S. Navya et. el.,"Comprehensive Women's Health Risk Assessment Platform", IEEE - Explore, 2ndIAIMLA - Conference, DOI: 10.1109/AIMLA59606.2024.10531341, ISBN: 979-8-3503-4923-8, PP. 1611 - 1619, 2024.

[4] Palvi Soni et. el.,"Machine Learning Applications in the Diagnosis of PCOS: A Systematic Review", IEEE - Explore, CYBERCOM - Conference, DOI:

10.1109/AIMLA59606.2024.10531341, ISBN: 979-8-3315-4235-1, PP. 1221 - 1232, 2024.
[5] Dr B.V. RamaKrishna et. el.,"Support Vector Machines for Early Detection of Chronic Diseases in Healthcare", Frontiers in Health Informatics, Vol.13, Issue-3, ISSN: 2676 - 7104, PP. 3792 - 3803, 2024.
[6] Yue Wu et. el., "Risk prediction model based on machine learning for predicting miscarriage among pregnant patients with immune abnormalities", Frontiers in Pharmacology, Vol. 15, Issue -7, ISSN: 1663-9812,DOI:10.3389/fphar.2024.1366529, PP. 134-145, 2024.
[7] Kimberley Kavanagh et. el., "Machine learning-enabled maternal risk assessment for women with pre-eclampsia (the PIERS-ML model): a modelling study", ELSEVIER, Vol. 6, Issue - 4, PP. 238 - 250, 2024.
[8] Alaa O Khadidos et. el., "Maternal Health Risk Detection: Advancing Midwifery with Artificial Intelligence", PubMed Central,Scientific Reports, Vol. 13, Issue - 7, DOI: 10.1038/s41598-024-71934-x, PP. 1-21, 2024.
[9] P.Arunkumar et. el.,"Application using Machine Learning to Promote Women's Personal Health", IEEE - Explore, 5thICSSIT - Conference, DOI: 10.1109/ICSSIT55814.2023.10061126, ISBN: 978-1-6654-7468-9, PP. 123 - 131, 2023.
[10] Priya Patil et. el., "AI-ML-Enabled Clinical Decisions for Pregnant Women and their Doctors with Special Attention to International Legal Aspects", IEEE-Explore, 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), ISBN: 2687-7759, DOI:10.1109/UPCON 59197.2023.10434549, PP. 1214-1221, 2023.
[11] Elena Korytnikova et. el., "Artificial Intelligence and Women's Health: Innovations, Challenges, and Ethical Considerations", Genisis Publications, DOI: 10.52793/ACMR.2023, Vol. 4, Issue - 3, PP. 23 - 29, 2023.
[12] Javeria Ali et. el., "Machine Learning Approaches for Prediction of Nutrition Deficiency among Women of Different Age Groups", IEEE-Explore, 3rd ICONICS - Conference, ISBN: 978-1-6654-6142-9, DOI: 10.1109/ICONICS56716.2022.10100340, PP. 112-121, 2022.
[13] Sharathkumar Anbu et. el., "Machine learning approach for predicting womens health risk", IEEE-Explore, ISBN: 978-1-5090-4560-0, DOI: 10.1109/ICACCS.2017.8014588, PP. 1121-1129, 2017.
[14] B.V. RamaKrishna et. el., "Mobile Cloud Diabetic Control System", IJCSMC, Vol. 6, Issue - 7, ISSN: 2320 -088X, PP. 121-127, 2017.