International Journal of Innovative Technology and Interdisciplinary Sciences
https://journals.tultech.eu/index.php/ijitis
ISSN: 2613-7305
Volume 9, Issue 1
DOI: https://doi.org/10.15157/IJITIS.2026.9.1.566-596
Received: 01.12.2025; Revised: 21.02.2026; Accepted: 22.03.2026

Research Article

Hybrid Crow Search and ICA Optimized Regression Random Forest Model for Precise Preterm Pregnancy Risk Prediction Assessment

Pritika Goel*, Sachin Ahuja

Department of Computer Science and Engineering, Chandigarh University, Mohali, India
***goelpritika320@gmail.com**

Abstract

This paper suggests a hybrid machine learning method, called Hybrid Crow-ICA with Regression Random Forest (HyCIRRF), to estimate the likelihood of having a preterm pregnancy based on maternal electronic health records (EHRs). It consists of a framework of Fast Independent Component Analysis (Fast-ICA) to extract features, CatBoost to rank the importance of features, and Crow Search Optimization (CSO) to select the best features. The last predictive ensemble is a hybrid of Regression Random Forest (RRF) and Logistic Regression (LR) with a soft-voting mechanism to enhance the productivity and interpretability. The model was tested on the Maternal Health Risk Dataset, and the model had an accuracy of 93%, a precision of 93%, a recall of 94, and an error rate of 7, which is the best balance between sensitivity and specificity. Moreover, the values of AUC greater than 0.95 represent great discrimination of low-, medium-, and high-risk groups. Compared to other baseline models, such as the Logistic Regression, Decision Tree, SVM, and XGBoost, HyCIRRF proved to be as accurate as those with explainability and computationally efficient simultaneously. These results indicate that this proposed model may be applied as a valid and understandable clinical decision support system in the early identification of high-risk pregnancies.

Keywords: Preterm Birth; Hybrid Model; Crow Search Optimization; Fast Independent Component Analysis; CatBoost; Logistic Regression; Regression Random Forest

INTRODUCTION

Preterm birth, characterized as childbirth occurring before 37 completed weeks of gestation, continues to be a major global public health concern due to its strong association with neonatal illness and mortality. The world health organization (WHO) reports that over 15 million babies are delivered prematurely annually, accounting for more than 10% of global births [1]. The preterm birth rate is projected to be 10.6 percent in the Indian setting, and it varies significantly by region because of the disparities in socioeconomic status and healthcare access [2]. Preterm birth risk prediction shows strong accuracy because the implementation of timely interventions can maximize the results of the work with expectant mothers and infants when the pregnancy is identified as being at risk at an early stage [3]. Historically, preterm birth risk has been analysed based on clinical feature measurements, with obstetric history, maternal age, cervical length, and presence of infections being included. Although these indicators provide informative information, their predictive power is limited. The cervical length may be used to predict preterm birth among certain populations, but predictive validity varies among most clinical settings. Similarly, maternal age and obstetric history are also correlated with increased risk but will not be reliable enough to predict preterm birth on a case-by-case basis [4]. The prediction of preterm birth has been notably improved with the emergence of modern ML techniques. ML techniques can be used to handle complex data sets, including electronic health records, ultrasound, and lab results, to reveal hidden trends and risk factors that can be missed with traditional techniques [5, 6]. SVM models are highly accurate, recalling and specific in predicting preterm birth. Likewise, XGBoost models are demonstrated to perform well at predicting the gestational week to delivery with a focus on the potential of ML to personalise risk assessment. The incorporation of biomarkers into predictive modeling has also enhanced diagnostic accuracy. Notably, C-reactive protein (CRP), interleukin-6 (IL-6), and interleukin-1β (IL-1β) are well-established in this context. Inflammatory markers that show a strong association with preterm birth risk [7]. Moreover, integrative multi-omics analyses, encompassing genomics, transcriptomics, proteomics, and metabolomics, are under research in order to identify new biomarkers and use them to build more predictive and sustainable predictive models [8].

In spite of these considerable advances, the standardisation of datasets and validation of predictive models in diverse populations still pose major issues and limit their generalizability and clinical utility. The combination of the various types of data used with genetic, environmental, and lifestyle data requires the use of complex analytical tools and raises significant ethical and privacy concerns [9]. Subsequent studies must focus on the development of a unified, multi-modal databank as well as the implementation of explainable AI methods to maintain the ability of the models to be read and accepted by medical practitioners. Integrating machine learning methodology with biomarker analysis is a potentially successful approach to risk prediction of preterm birth. It offers the potential to reduce neonatal morbidity and mortality and enhances maternal health outcomes by identifying pregnancies at high risk, which is likely to result in timely clinical interventions.

Although the state of maternal healthcare has improved, the proper identification of the women who are at risk of preterm pregnancy has been a significant clinical issue because the interactions between physiological, clinical, and lifestyle factors are complex and nonlinear. Conventional methods of screening usually use a few indicators and clinical judgment, which might fail to capture hidden patterns that lead to the development of early pregnancy complications. Current machine learning models have demonstrated the prospects, but most of them are still associated with redundancy of features, poor optimization of features, lower generalizability, and diminished interpretability, rendering them hard to apply in practice in clinical settings. Hence, a predictive model with reliable, optimised, and interpretable characteristics is necessary to analyze critical maternal health parameters and help detect high-risk pregnancies to enhance preventive care and maternal-fetal outcomes.

This research introduces the HyCIRRF model, which is a novel approach for preterm pregnancy risk prediction using Crow Search Optimization (CSO) and Random Forests and Fast Independent Component Analysis (FastICA). The major contributions and innovations of this study are as follows:

- Integration of FastICA for Feature Extraction Unlike traditional methods that utilise basic feature selection methods, FastICA is used for dimensionality reduction and for extracting independent components. This approach helps to improve the quality of features used for model training, which can help to improve the predictive accuracy of the model by reducing multi collinearity and redundancy in the data.
- Use of Crow Search Optimization (CSO): CSO is used for feature selection, which is a computationally efficient and faster convergence method than traditional methods such as Genetic Algorithms (GA). By taking advantage of the intelligent search behavior of crows, CSO optimizes subsets of features, so that the model uses the most informative features to make accurate predictions.
- Ensemble Learning for Robustness: The proposed model includes ensemble learning, where the strengths of various classifiers are combined, resulting in better generalization and stability. This leads to a more robust model that can handle complex, real-world data, especially in the case of medical risk prediction.

The combination of FastICA, CSO, and ensemble learning is a novel approach to predictive modelling, which has shown superior performance in the preterm pregnancy risk prediction task compared to traditional methods.

LITERATURE REVIEW

The development of preterm birth prediction has shifted the old-fashioned risk scoring to the data-driven approach, which incorporates clinical measurements, electronic health records, physiological signals, and imaging. Recent research demonstrates that machine learning has the potential to enhance early risk stratification, but performance, generalizability, and clinical interpretability are the main issues. The section is a summary of representative work in three directions: (i) general ML methods on preterm birth prediction, (ii) optimization-based and hybrid ML methods in maternal health, and (iii) explainable AI methods to enhance clinical trust.

Machine Learning Approaches for Preterm Birth Prediction

Authors in [10] conducted a survey of machine learning methods to predict preterm birth with a variety of data modalities, such as electrohysterogram (EHG) signals, electronic health records, and ultrasound-based measurements. The research noted that existing clinical screening has the potential to miss a significant number of cases, which is partly because of the complexity of risk interactions and inconsistency in evaluation. The review highlighted that data-based models can be used as supplements to clinical assessment, but the implementation of the models can be hampered by variability of the data, inconsistency in measurements, and difficulties in model validation across the different settings.

Researchers at [11] have developed a systematic review dedicated to the preterm birth prediction with the help of health record data and machine learning. According to their analysis, the study designs differ greatly in the number of cohorts and timeframes, with hundreds of cases to very big multi-year datasets. Study noted that model performance is usually reported in terms of accuracy, sensitivity, specificity, and ROC-AUC and also found that there are common weaknesses in methodological reporting, especially in relation to feature selection decisions, software implementation decisions, and data handling decisions. The review found that more robust justification of modelling choices and more standardised assessment protocols are required in order to have reliable clinical adoption.

Authors in [12] created a machine learning model that predicts preterm birth in the rural India setting. Study focused on the selection of features with the help of an entropy-based concept to select informative maternal indicators out of obstetrical data. Several classifiers were tested and the Support Vector Machine reported the best performance among the tested models. The paper proved that location-specific modelling may provide useful results but the techniques depend on the quality and representativeness of the gathered clinical data and the chosen feature space.

In general, the literature suggests that machine learning can offer potentially useful predictive value, particularly when used on structured clinical data or physiological data. Simultaneously, numerous studies vary in datasets, validation configurations, and reporting detail, making it hard to directly compare them and restricting their reproducibility.

Optimization-Based and Hybrid ML Models in Maternal Health

To improve the quality of prediction in maternal health tasks, optimization strategies are becoming more popular, typically through the tuning of model parameters, model feature selection, or learning stability in the presence of class imbalance.

The authors in [13] suggested a parameter optimization method based on the Mud Ring Algorithm (MRA) to improve the prediction of maternal health risks. Their work proposed a two-step approach that included optimization of Support Vector Machine parameters with the help of MRA followed by imbalance mitigation with an oversampling technique, and finally the training of optimised classifier, e.g., SVM, Random Forest, and KNN. Study indicated that the integration of optimization and imbalance-conscious preprocessing can enhance predictive quality on clinical risk data, as their experiments have shown significant improvements in multiple measures of evaluation.

Researchers at [14] introduced a machine learning architecture that is hyperparameter-optimised to identify foetal genetic disease at the first trimester of pregnancy, based on first-trimester screening. Study used class balancing by synthetically generating samples and feature engineering to enhance interpretability and robustness in their pipeline. Among several other models that were tested, CatBoost reported the best performance. This study illustrates the importance of optimising predictive models and designing clinically significant feature interactions as a way of improving the results of early screening.

Authors in [15] researched the concept of biologically inspired optimization of training neural networks in e-health pregnancy care environment. Study employed Particle Swarm Optimization in their work to optimise the structure and parameters of neural networks and found that study improved the precision and true positive rate, and reduced the false positive rate and validation error. The research helps to substantiate the general opinion that nature-based optimization may be useful in clinical decision support systems by enhancing performance without making the model too complex.

To conclude, optimization-based models and hybrid frameworks are typically created to overcome the typical limitations in maternal health modelling, such as parameter sensitivity, redundant features, computational cost, and imbalanced distribution of classes.

Explainable AI and Clinical Interpretability in Pregnancy Risk Prediction

The consequences of misclassification are high and the clinical reasoning needs to be clear; thus, interpretability has become a significant requirement in the context of pregnancy-related decision support.

Authors in [16] examined the prediction of birth weight based on early and mid-pregnancy antenatal markers and combined machine learning with explainable AI methods, including SHAP, LIME, and Anchor. The paper has highlighted that explainability may enhance clinical confidence by discovering important predictors and giving justifications behind individual predictions despite the low accuracy due to the size of the data set and population characteristics.

Authors in [17] created a clinical decision support system that is explainable and based on ML to predict gestational diabetes mellitus. Data preparation, class balancing through SMOTE, feature selection, and model explanation through SHAP were used in the study to enhance trust and acceptability. Their findings revealed that explainability can be incorporated with model selection and tuning so that models can be used to support focused interventions as opposed to general screening policies.

A prospective cohort was employed by [18] to predict preterm birth with explainable machine learning and SHAP (and LIME to explain each instance in particular) to determine key risk factors. Their work emphasised the fact that interpretability is not just useful in feature ranking, but also in patient-specific risk reasoning, which may be used to support clinical decision-making and communication.

In a study of [19], the researchers compared the traditional ML-based models with deep learning (transformer-based) models on a large retrospective dataset to predict preterm birth. Transformer model performed better than the logistic regression, random forest and SVM during their assessment. The research indicates the increasing popularity of deep learning in large-scale prediction problems, as well as creates practical concerns regarding the size of data requirements and interpretability in comparison to less complex models.

Together, these papers suggest that explainable AI techniques can help transform predictive models into more practical use in healthcare, because studies can be used to attribute features, provide personalized reasoning, and enhance clinical confidence. Table 1 shows the Comparative SOTA Analysis of the existing work.

Table 1. Comparative Summary of Representative SOTA Studies Related to Pregnancy Risk/Preterm Birth Prediction

Ref. No.	Study Focus	Dataset Size	ML Models Used	Best Model	Outcome (Quantitative Results)
[11]	Preterm Birth (Systematic Review)	274 – 1,400,000 (across studies)	Multiple ML models	Varies	13 eligible studies identified; data span 1–11 years; model performance commonly reported using accuracy, sensitivity, specificity, and AUC (exact values varied across studies)
[12]	Preterm Birth Prediction (Rural India)	NR	DT, LR, SVM	SVM	Achieved 90.9% accuracy, outperforming Decision Tree and Logistic Regression
[13]	Maternal Health Risk Prediction	13 datasets	SVM, RF, KNN with MRA	MRA-RF	Prediction accuracy increased by 11.8% (SVM), 9.11% (RF), and 17.08% (KNN) compared to non-optimized models
[14]	Down Syndrome Risk Prediction	959 pregnancies	8 ML classifiers	CatBoost	Achieved 97.39% accuracy with a 2.62% false-positive rate, outperforming

					XGBoost and other classifiers
[15]	High-Risk Pregnancy CDSS	NR	MLP, PSO-MLP	PSO-MLP	Improved precision by 26.4%, increased TPR by 14.9%, reduced FPR by 35.4%, and improved AUC by 10.2% compared to baseline MLP
[16]	Low Birth Weight Prediction	237 pregnancies	AdaBoost, Ensemble ML	AdaBoost	Achieved 77% accuracy, 73% precision, 77% recall, and 72% F1-score
[17]	Gestational Diabetes Mellitus Prediction	NR (PEARS study)	5 ML models	Use-case dependent	Best model achieved AUC-ROC = 0.792 (theoretical model); screening models achieved AUC-ROC ≈ 0.66
[18]	Preterm Birth with Explainability (XAI)	3,509 pregnancies	6 ML models	XGBoost	Achieved AUC = 0.735 (parous) and 0.723 (nulliparous); overall PTB incidence 11.23%
[19]	Preterm Birth Prediction (Deep Learning)	30,965 births	LR, RF, SVM, Transformer	Transformer	Achieved AUC = 79.20%, accuracy = 72.61%, sensitivity = 73.67%, specificity = 72.48%, NPV = 94.95%

Research Gaps

Based on the studies reviewed, the gaps that become critical directions in structured-data pregnancy risk prediction are as follows:

- Poor consistency of methodological description and reproducibility: Most of the studies report good results but lack transparency of hyperparameters, feature selection process, and evaluation procedures, making it hard to replicate them.
- Redundancy and correlation of features in standard clinical measurements: Physiological variables (e.g., blood pressure components) are usually multicollinear, and unless dimensionality reduction or transformations based on independence is used, they can be unstable.
- Tuning is more commonly considered to be optimization, rather than pipeline-level selection: Other works optimize hyperparameters, although there are fewer works that combine metaheuristic optimization to choose small, informative feature subsets with similar evaluation criteria.

- Interpretability is often an afterthought to model training, and is not part of the pipeline: The concept of explainability is often used as a post-hoc procedure, but in practise, it is better to think of interpretability during the pipeline (importance of its features, decision logic, and model selection).
- The model performance is frequently data specific: Research often uses domain/cohort-specific data, and results can shift when there is a change in feature distributions. This encourages prudent pipeline design which focuses on stability, bias control and sound evaluation.

Problem Statement

Preterm birth is one of the leading causes of neonatal morbidity and mortality and the clinical intervention in high-risk pregnancies should be identified early to provide timely clinical intervention. Despite the promising results of recent machine learning methods in predicting preterm birth based on electronic health records, clinical variables and biomarkers, there are still a number of important limitations [21-26]. Current models are prone to feature redundancy and multicollinearity, are based on single or partial optimization methods and are often not directly interpretable, making them less acceptable in clinical practice.

In addition, most of the top-performing methods rely on complicated deep learning models or large amounts of multimodal data, which are computationally intensive and cannot be implemented in typical healthcare environments. The optimization methods are usually limited to hyperparameter optimization or imbalance management, and the joint issues of feature independence, optimal feature selection, predictive stability, and explainability are not considered on a single framework.

Thus, there is a definite necessity of a robust, optimised, and interpretable machine learning framework capable of using organised maternal health data to properly predict the risk of preterm pregnancy, reduce redundancy of features, ensure clinical transparency, and enable its practical implementation as a decision-support tool in maternal medical care.

The main hypothesis of our research work is as follows:

- H1: The application of Fast Independent Component Analysis (Fast-ICA) leads to a reduction of multi collinearity between maternal clinical features and an enhancement of the stability of preterm pregnancy risk prediction.
- H2: Crow Search Optimization (CSO) based on the feature importance determined by CatBoost selects an optimal set of maternal health features which improves the classification accuracy compared to non-optimised feature sets.
- H3: The hybrid soft-voting ensemble of Regression Random Forest and Logistic Regression demonstrates competitive, stable, and clinically reliable predictive performance compared to individual and traditional machine learning classifiers.

- H4: The proposed HyCIRRF framework is able to discriminate between low, mid, and high-risk classification of pregnancy with high sensitivity and overall classification accuracy.

PROPOSED METHODOLOGY: HyCIRRF FRAMEWORK

This part gives the full methodological pipeline of the proposed Hybrid Crow Search and ICA optimised Regression Random Forest (HyCIRRF) model. The framework will be capable of achieving the accurate, stable and interpretable prediction of the preterm pregnancy risk by incorporating dimensionality reduction, feature ranking, metaheuristic optimization and ensemble learning in a single architecture.

Overall System Architecture

The HyCIRRF model is sequential and modular in nature and converts raw data on maternal health into clinically significant risk forecasts, see Figure 1.

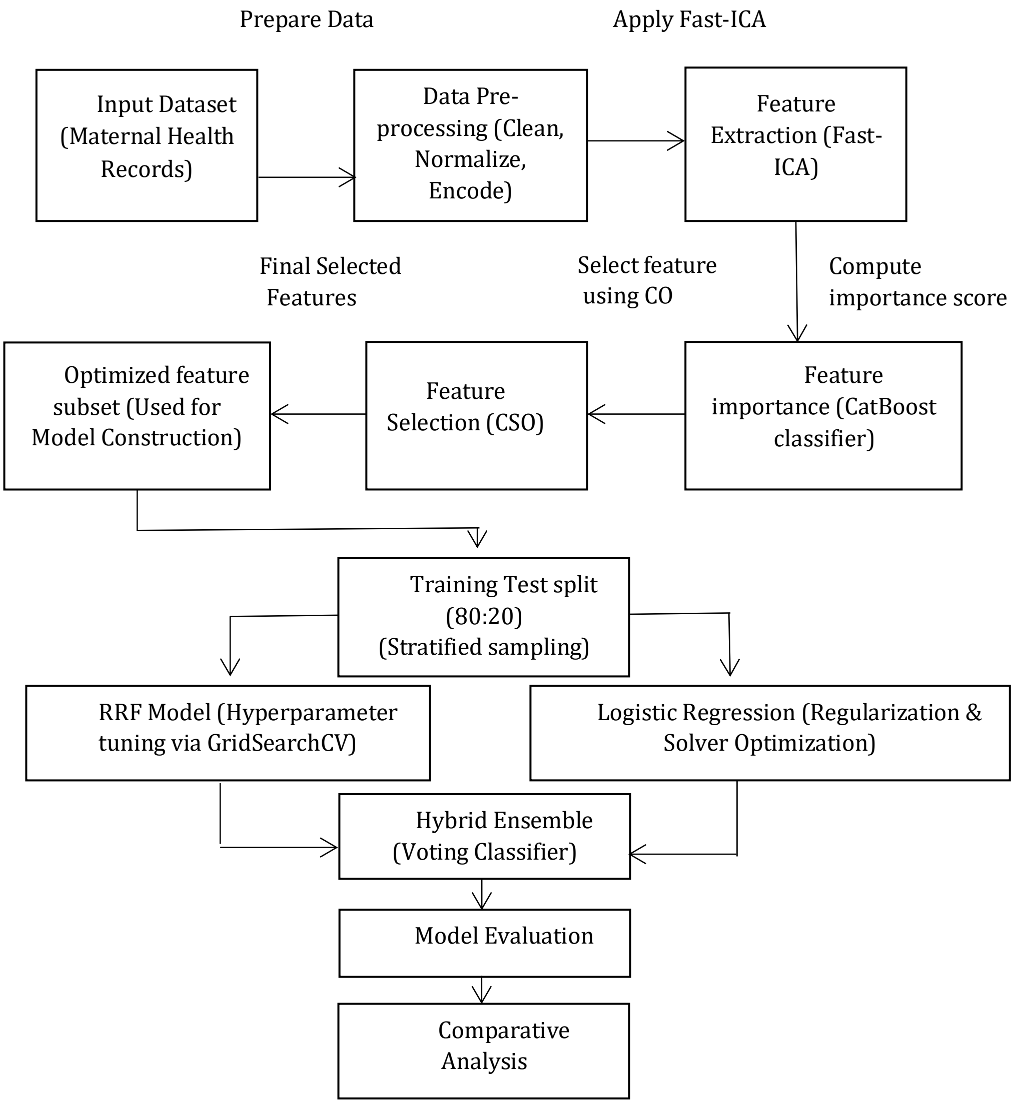

Figure 1. Flowchart of the Proposed Work

The steps involved in the workflow are data preprocessing, multistage feature engineering, and optimised ensemble classification.

First, raw electronic health record (EHR) data are purged, normalised and divided into training and testing subsets. Fast Independent Component Analysis (Fast-ICA) is then used to extract features and eliminate redundancy as well as identify statistically independent latent components. CatBoost is used to determine the importance of original clinical features relative to each other in order to guarantee interpretability. These ranked features are then optimised using Crow Search Optimization (CSO) that finds the most informative subset by maximising predictive performance.

The optimised features are then given to a hybrid ensemble classifier, which is composed of Regression Random Forest (RRF) and Logistic Regression (LR). A soft-voting mechanism is a probabilistic mechanism that takes the outputs of both models to come up with the final risk classification (Low, Mid, or High). This multilevel architecture provides strength, less overfitting, and improved explainability. Below figure 1 describe the complete End to End pipeline of the proposed work.

Dataset Description

The data applied in this study is the Maternal Health Risk Dataset, which is publicly accessible on the Kaggle repository [20]. This data is commonly used in maternal health analytics and as a reference in the prediction of pregnancy risks by machine learning.

The data set includes 1,014 patient records, where each of the records describes the clinical profile of a pregnant woman. It has seven medically relevant attributes, which are maternal age, systolic blood pressure, diastolic blood pressure, blood glucose level, body temperature, heart rate, and a risk label that is categorical. The target variable is used to classify the risk of pregnancy as Low, Mid, and High.

The completeness of this dataset is also one of its strongest features, since the dataset is not missing any values, and it is possible to train the models without imputation bias. Also, the characteristics are related to periodically measure clinical indicators, and the data is applicable to real-world decision support modelling. Below Table 2 shows Characteristics and Feature Description of Maternal Health Risk Dataset.

Table 2. Maternal Health Risk Dataset: Characteristic and Feature Description

Category	**Attribute / Feature**	**Description**	**Type / Range**
Dataset Overview	Data Type	Structured tabular clinical data	—
	Number of Instances	1,014 maternal records	—
	Data Source	Public benchmark dataset (Kaggle)	—
	Missing Values	None	Complete dataset
	Suitability for ML	Clean, normalized, structured	High

Target Variable	RiskLevel	Pregnancy risk classification	Categorical (Low, Mid, High)
Clinical Features	Age	Age of pregnant woman	Numerical (10–70 years)
	SystolicBP	Systolic blood pressure	Numerical (90–180 mmHg)
	DiastolicBP	Diastolic blood pressure	Numerical (50–120 mmHg)
	BS	Blood glucose level	Numerical (6–20 mmol/L)
	BodyTemp	Body temperature	Numerical (95–105 °F)
	HeartRate	Heart rate	Numerical (60–100 bpm)

Data Preprocessing and Class Imbalance Handling

Some preprocessing operations were used in order to guarantee the quality of data and stability of the model. Missing values and duplicate records were checked in the dataset; none were detected, which proves the completeness of the dataset. The nominal target variable of the levels of pregnancy risk was coded in numbers to enable it to be compatible with machine learning algorithms.

Z-score normalisation was used to standardise all numerical characteristics to a common scale, which enhances convergence and eliminates the dominance of high-magnitude characteristics. An 80:20 stratified split was then used to split the dataset into training and testing sets, making sure that the original class distribution was maintained within the subsets.

Though the dataset has moderate class imbalance, stratified sampling was used to minimise bias in the training process. The ensemble-based design of HyCIRRF also reduces the imbalance-related problem by integrating the linear and non-linear decision boundary.

Feature Extraction using Fast Independent Component Analysis (Fast-ICA)

The Fast-ICA was used to convert correlated clinical variables into statistically independent ones. In contrast to the classical dimensionality reduction techniques which are based on maximising the variance, Fast-ICA is aimed at maximising non-Gaussian, which allows it to isolate independent sources of information.

Let the normalized maternal health dataset be represented as equation (1):

$$\mathrm{X} \in R^{n\times \mathrm{d}} \tag{1}$$

Where n is the number of samples and d correspond to the number of clinical features.

Whitening Transformation is expressed via equation (2).

$$\mathrm{Z} = D^{-1/2}E^{T}(X - E[X]) \tag{2}$$

Where E and D are eigenvectors and eigenvalues of the covariance matrix.

Fast ICA maximizes non-Gaussianity, and is expressed via equation (3)

$$w_{max} \mathrm{J}(\mathrm{w}^T \mathrm{Z}) = \mathrm{E}[\mathrm{G}(\mathrm{w}^T \mathrm{Z})] - \mathrm{E}[\mathrm{G}(\mathrm{v})] \tag{3}$$

Where $G(\cdot)$ is a contrast function (e.g., log-cosh), $v \sim N(0,1)$.

Fixed-point update rule has been found via equation (4)

$$w^{t+1} = E\left[Zg\left(w^{(t)T} Z\right)\right] - E[g'(w^{(t)T} Z]W^{(t)} \tag{4}$$

This produces statistically independent components through equation (5):

$$S = wz \tag{5}$$

In the present work, the original feature space was broken down into five independent components by a fixed-point iteration scheme. This conversion minimises multicollinearity in the physiological measures like systolic and diastolic blood pressure and still preserves information of clinical interest. The derived elements are small-scale depictions of maternal health trends and enhance the efficiency of calculations in later phases.

Feature Importance Ranking using CatBoost

After the extraction of features, the contribution of each of the original clinical attributes to the prediction of pregnancy risks was estimated with CatBoost. CatBoost can efficiently work with structured healthcare data because it is robust and has an inbuilt ability to address the interactions between features.

CatBoost minimizes the following loss;

$$\mathrm{L} = \sum_{i=1}^{n} l(y_i \mathrm{y}^{\wedge}{}_i) \tag{6}$$

Feature importance is calculated using prediction loss change:

$$FI_j = E[L(X) - L\left(X_{\setminus j}\right)] \tag{7}$$

Where, $X_{\setminus j}$ denotes the datset with feature j removed, and Higher FI_j indicated stronger contribution to risk prediction.

The pre-processed dataset was then trained on the model, and the scores of importance of any feature were obtained by the effect of the given variable on the predictive loss of the model. This action brings transparency since it brings to the fore the risk indicators that are predominant, including blood pressure and the level of blood glucose. The scores of importance ranking serve as a control mechanism to the optimization phase, making the process of feature selection data-driven and understandable.

Feature Selection via Crow Search Optimization (CSO)

To establish the best subset of features, which is the least redundant and maximises classification accuracy, Crow Search Optimization was used. CSO is a metaheuristic population-based algorithm that is based on intelligent food-hiding and retrieval behaviour of crows.

Each crow is a candidate feature subset in the form of a binary vector in the optimization process. Fitness The fitness of each subset is measured based on the mean classification accuracy of a 5-fold cross-validated Random Forest classifier. Crows search the feature space by updating their positions iteratively and learning through memory, and narrow down to the most informative subset.

Each crow represents a binary feature subset, see equation (8):

$$P_i = [p_{i1}, p_{i2}, p_{i3}, \ldots \ldots p_{id}], \quad p_{ij} \in \{0,1\} \tag{8}$$

Position update rule

$$p_i^{t+1} = \begin{cases} P_i^t + r_i . fl_i . (m_j^t - P_i^t), & if\ r_j \geq AP \\ random\ position, & otherwise \end{cases} \tag{9}$$

Were, $m_j -$ memory of crow j, $fl = mf$light length, $AP -$ awareness probability

Fitness Function

$$Fitness = \frac{1}{K} \sum_{k=1}^{K} Accuracy_{RF}^{(k)} \tag{10}$$

This dimensionality reduction and predictive strength optimization process is effective in balancing dimensionality reduction and predictive strength, which results in better generalisation and less computational overhead.

Hybrid Ensemble Learning

HyCIRRF framework uses hybrid ensemble learning approach which combines tree-based model which is non-linear with a linear probabilistic classifier. Complementary nature of these learners makes it possible to have strong prediction, better generalisation, and better interpretability. The ensemble decision is generated with the help of soft-voting mechanism which sums up probabilistic outputs. Table 3 shows Hybrid Ensemble Learning Components of the HyCIRRF Framework.

Regression Random Forest is employed in the proposed ensemble architecture to provide the non-linear decision-making capacity in the ensemble that is effective in capturing physiological indicator interactions, and Logistic Regression provides the linear interpretability in the form of coefficients that provide probability estimates. The soft-voting approach calculates the average probability of each class in both models and the one with the highest sum of probability is chosen as the final output.

This ensemble design makes sure that none of the modeling paradigms takes over predictions, thus enhancing robustness and clinical credibility. The probabilistic fusion also reduces the misclassification risk among borderline cases which is extremely important in the maternal health risk assessment.

Table 3. Hybrid Ensemble Learning Components of the HyCIRRF Framework

Component	Algorithm	Role in Framework	Key Characteristics	Contribution to Prediction	Mathematical Formulation
Base Learner 1	Regression Random Forest (RRF)	Captures non-linear relationships and complex feature interactions	Ensemble of decision trees trained on bootstrapped samples with randomized feature selection	Provides high predictive power, robustness to noise, and resistance to overfitting	$yRF = \frac{1}{T}\sum_{t=1}^{T} f_t(x)$
Base Learner 2	Logistic Regression (LR)	Models linear associations between features and risk outcome	Regularized probabilistic classifier with interpretable coefficients	Enhances transparency and supports clinical interpretability	$P_{LR}(y=1\|x) = \frac{1}{1+e^{-(w^T x+b)}}$
Ensemble Strategy	Soft-Voting Mechanism	Combines probabilistic outputs of RRF and LR	Averages class probability distributions rather than hard labels	Reduces variance, balances bias, and stabilizes predictions	$P_c = \frac{1}{2}(P_{RF}(c\|x) + P_{LR}(c\|x))$
Final Output	Hybrid Prediction	Generates final pregnancy risk classification	Risk levels: Low, Mid, High	Improves reliability and decision confidence in clinical use	$y = argmax_c P_c$

Algorithmic Pseudo Code and Hyperparameter Settings

Algorithm: HyCIRRF Framework for Preterm Pregnancy Risk Prediction.
Input: Maternal Health Dataset
Output: Pregnancy Risk Class (Low / Mid / High)

Algorithm: HyCIRRF Framework

Load dataset and perform feature normalization

2. Encode categorical target variable

3. Split dataset into training and testing sets using stratified sampling

4. Apply Fast Independent Component Analysis (Fast-ICA) on training data

5. Train CatBoost classifier and compute feature importance scores

6. Initialize Crow Search Optimization (CSO) population with random binary feature subsets

7. Evaluate fitness of each subset using 5-fold cross-validated Random Forest accuracy

8. Update crow positions and memory until convergence criteria are met

9. Select the optimal feature subset based on maximum fitness
10. Train Regression Random Forest (RRF) and Logistic Regression (LR) models using selected features
11. Combine class probability outputs using soft-voting strategy
12. Output final pregnancy risk classification.

Table 4 shows the hyper parameter used in the proposed work.

Table 4. Key Hyperparameter Settings for the HyCIRRF Framework

Model / Technique	**Hyperparameter**	**Value / Setting**
Fast Independent Component Analysis (Fast-ICA)	Number of components	5
	Convergence tolerance	1×10−41 \times 10^{-4}1×10−4
CatBoost Classifier	Number of iterations	1000
	Tree depth	6
	Learning rate	0.03
Crow Search Optimization (CSO)	Population size	20
	Maximum iterations	100
	Awareness probability	0.1
Regression Random Forest (RRF)	Number of trees	100
	Maximum tree depth	Optimized using GridSearch
Logistic Regression (LR)	Regularization type	L2
	Optimization solver	liblinear

Statistical Validation and Evaluation

The cross-validation of the HyCIRRF model was repeated five times with 10 to assess the robustness and generalizability of the model. The approach guarantees that a single train-test split does not bias the performance measures and gives a good idea of how the model will perform when applied to different subsets of the data.

Statistical significance tests were done to confirm the performance difference between the proposed HyCIRRF model and the baseline models.

The statistical significance of the differences in the accuracy of the models was evaluated using the paired t-test and the Wilcoxon signed-rank test. The results of both tests had p-values of more than 0.05, which shows that the difference in accuracy between the two tests was not significant.

Additionally, the test of McNemar was carried out with the result of p-value 1.000, which indicates that the overall performance is different, but the structure of misclassification is similar at the instance level in both models.

Calibration

To measure the calibration of the HyCIRRF model, the Brier score was employed because it is one of the most important measures to assess the probabilistic predictions in clinical practise. The Brier score of the model was 0.0378, lower than the value of 0.05, which means that the model is well calibrated and the predicted probabilities are close to the observed ones.

A calibration curve was also drawn, which proved the valid probabilistic estimation of the model, which guarantees that the predicted frequencies are equal to the observed frequencies. This is especially important in clinical decision-making, where sound probability estimates may avert overconfident false identifications and enhance the evaluation of risk.

EXPERIMENTAL SETUP

In this section, the experimental setup that was used to determine the efficacy of the proposed HyCIRRF framework is described. It describes the computational setting, validation plan, performance measures and comparative baselines to determine the predictive performance and clinical reliability of the model.

Experimental Environment and Tools

All experiments were conducted in a controlled computational environment to ensure reproducibility and consistency of results as in Table 5.

Table 5. Experimental Environment Configuration

Component	Specification / Tool Used	Purpose
Programming Language	Python	Model implementation and experimentation
Data Processing Libraries	NumPy, Pandas	Data cleaning, transformation, and analysis
Machine Learning Framework	Scikit-learn	Model training, validation, and evaluation
Feature Importance Tool	CatBoost	Feature ranking and importance estimation
Optimization Technique	Crow Search Optimization (CSO)	Metaheuristic feature selection
Visualization Tools	Matplotlib, Seaborn	Graphical interpretation of results
Execution Platform	Standard desktop computing system	Ensures reproducibility and stability

The choice of these tools ensures transparency, reproducibility, and alignment with commonly accepted practices in medical machine learning research.

Cross-Validation Strategy and Data Splits

In order to provide the estimation of performance without any bias, the dataset was split into 80:20 train and test sets in a stratified manner, maintaining the same distribution of pregnancy risk classes in the subsets. The moderate imbalance in the number of classes in the dataset necessitated stratification.

In the training set, 5-fold cross-validation was used during feature selection with the Crow Search Optimization, to measure the fitness of candidate feature subsets. This internal validation method minimizes the chances of overfitting and facilitates sound feature selection.

To train the hybrid ensemble on the training data and to evaluate the performance metrics on the unseen test set, the optimised feature subset was used to train the hybrid ensemble on the training data. The division between optimization and evaluation stages is a guarantee of methodological rigour and equitable evaluation.

Evaluation Metrics and Clinical Relevance

The evaluation metrics were used to measure model performance in a manner that included the statistical accuracy and clinical significance. Table 6 shows the metrics that will be chosen are:

Table 6. Evaluation Metrics with Description

Metric	Description	Mathematical Formula
Accuracy	Overall correctness of predictions across all pregnancy risk classes	$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$
Precision	Proportion of correctly predicted high-risk pregnancies among all predicted high-risk cases	$Precision = \frac{TP}{TP + FP}$
Recall (Sensitivity)	Ability of the model to correctly identify true high-risk pregnancies	$Recall = \frac{TP}{TP + FN}$
F1-Score	Harmonic mean of precision and recall, effective for imbalanced datasets	$F1 = \frac{2 * Precision * Recall}{Precision + Recall}$
Error Rate	Proportion of misclassified instances	$Error\ Rate = \frac{FP + FN}{TP + TN + FP + FN}$

Moreover, the prediction behaviour at the class level was also analysed through the confusion matrix to determine the misclassification patterns between the neighbouring risk levels. The discriminative ability of the model against each risk class was calculated by means of a one-versus-rest method to obtain Receiver Operating Characteristic (ROC) curves and Area Under the Curve (AUC) values.

Clinically, recall and AUC received specific significance because the inability to identify high-risk pregnancies may result in serious maternal and neonatal complications.

Baseline Models and Comparative Protocol

In order to confirm the performance of the proposed HyCIRRF framework, the performance was compared to various popular machine learning classifiers that are regularly used in the risk prediction of preterm pregnancies. The chosen baseline models are:

- Logistic Regression (LR): It is a high interpretable linear probabilistic baseline.
- Decision Tree (DT): This is a rule-based classifier, which offers clear decision boundaries.
- Support Vector machine (SVM): It is useful in dealing with non-linear decision surfaces with small datasets.
- Extreme Gradient Boosting (XGBoost): It is an effective ensemble algorithm that has a high predictive accuracy.

The proposed model was trained on the same pre-processed data, feature sets and train test splits as all the baseline models to guarantee fairness in their comparison. Standard optimization procedures were used to select hyperparameters of each baseline where feasible.

Comparisons of the performance were done with the same evaluation metrics and visualisation. This protocol will make sure that the observed changes in performance can be attributed to the suggested hybrid optimization and ensemble strategy and not experimental bias.

RESULTS

The results of the experiment related to the Hybrid Crow–ICA with Regression Random Forest (HyCIRRF) model are presented in this section and were developed to predict the risk of preterm pregnancy. The model integrates Fast Independent Component Analysis (Fast-ICA) for feature extraction, CatBoost for feature importance evaluation, Crow Search Optimization (CSO) for feature selection, and a soft-voting ensemble of Logistic Regression and Random Forest for final classification. Each stage of the framework is evaluated through structured quantitative and visual analysis.

Independent Component Analysis (Fast-ICA)

The maternal health dataset was subjected to fast-ICA to eliminate redundant information and to determine statistically independent latent components. This transformation was useful in capturing the most informative and non-overlapping patterns of the clinical variables (such as age, blood pressure, glucose, and heart rate). The feature extraction and selection process flow is shown in Figure 2, which emphasizes the sequential application of the Fast-ICA, CatBoost, and CSO algorithms.

Afterward, feature importance analysis and optimization were performed using the five independent components that were produced.

The ability to reduce the correlated physiological variables into independent components led to stable classification behavior and contributed towards high discriminative performance with ROC--AUC values greater than 0.95 across all pregnancy risk classes, thus validating Hypothesis H1.

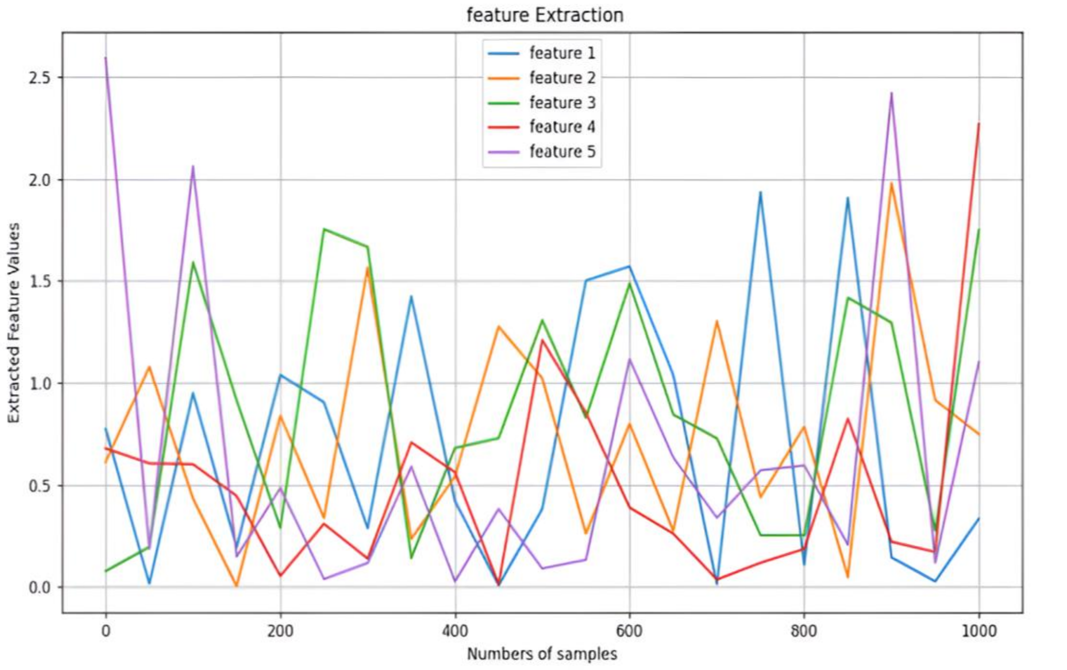

Figure 2. Feature Engineering and Selection Process

Multi collinearity Analysis and ICA Justification

To further justify the application of Fast-ICA, multi collinearity diagnostics were performed using Variance Inflation Factor (VIF) in Table 7.

Table 7. Variance Inflation Factor (VIF) values

Feature	VIF
Age	8.32
SystolicBP	105.68
DiastolicBP	85.31
BS	11.67
BodyTemp	126.77
HeartRate	89.54

VIFs that are greater than 10 represent acute multicollinearity. Some of the predictors have very high VIF values (>80), which confirms a high inter-feature linear dependency. The heatmap of correlation represented in Figure 3 also indicates that physiological variables are strongly associated.

Therefore, Fast-ICA was applied not for dimensionality reduction but for de correlation of latent physiological signals and stabilization of ensemble learning, directly addressing methodological justification concerns.

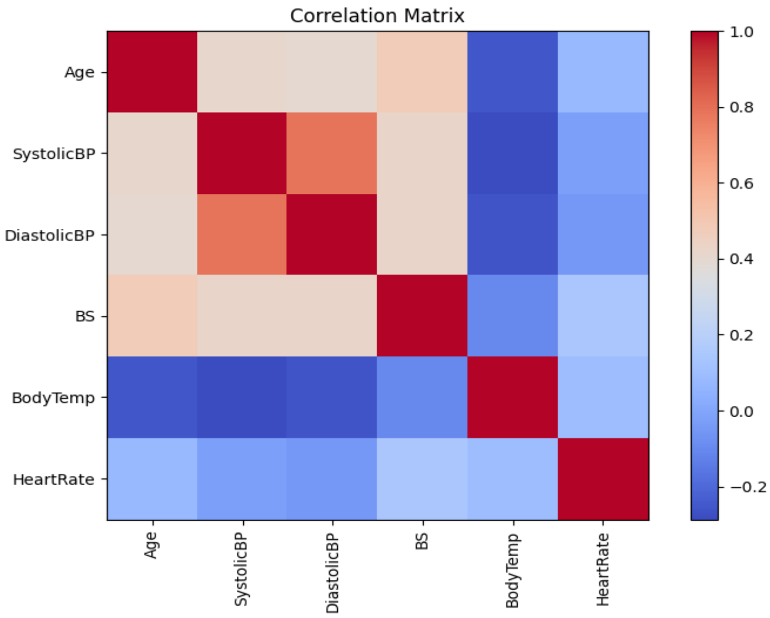

Figure 3. Correlation Heatmap of Maternal Clinical Variables

Feature Importance Using CatBoost

A CatBoostClassifier was trained on the processed data after ICA transformation in order to measure each variable's contribution to the prediction of pregnancy risk. Features like systolic blood pressure and blood glucose level were found to have the most predictive value in Figure 4, a horizontal bar plot of feature relevance. This phase serves as a link between the extraction of raw features and the later metaheuristic feature selection procedure.

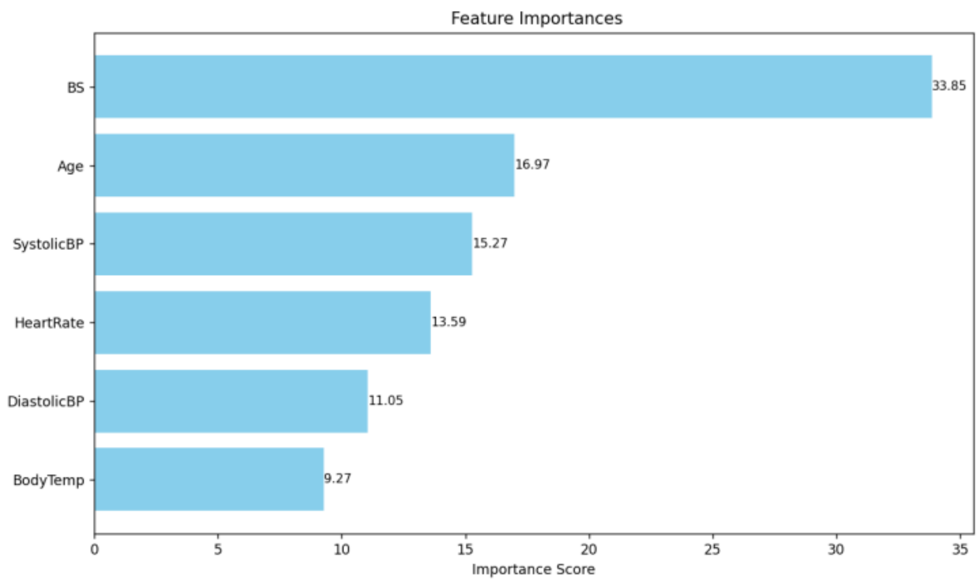

Figure 4. Feature Importance Scores derived from CatBoost

Feature Optimization Using Crow Search Optimization (CSO)

Figure 5 displays the convergence curve of CSO iterations and confirms the stability of the optimization process and a gradual increase in fitness over the course of the iterations.

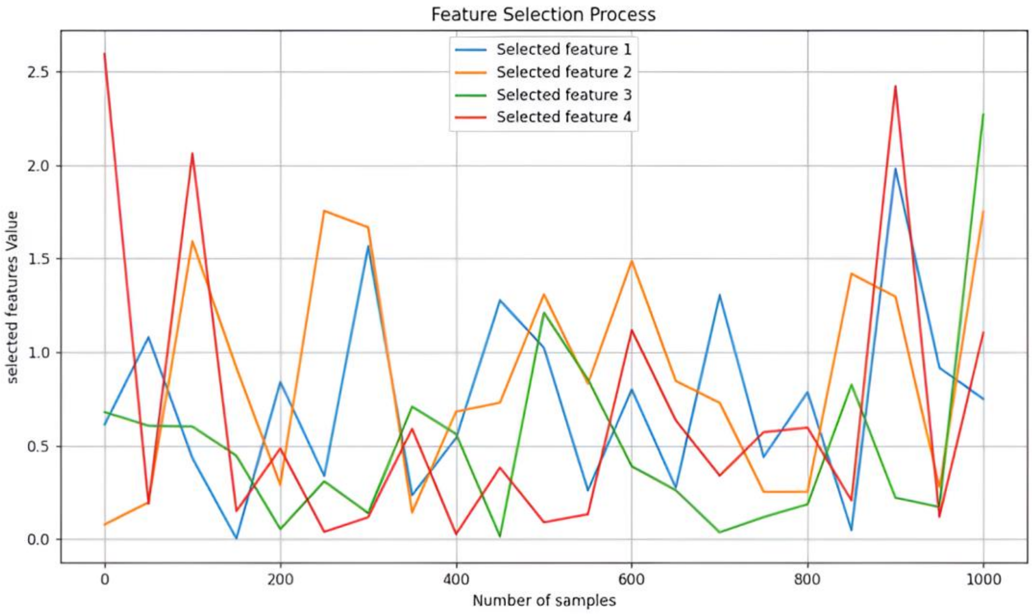

Figure 5. CSO Optimization Convergence and Selected Feature Subset

Using a Random Forest-based fitness function that was assessed using 5-fold cross-validation, the CSO method maximized the subset of characteristics by mimicking crows' clever search activity. To find the most important features that would balance the accuracy of the forecasting and the efficiency of the computing, iterative optimization was used.

The CatBoost-guided CSO process converged to an optimal feature subset that maximized cross-validated accuracy and reduced the overall error rate to 7%, supporting Hypothesis H2 that optimization-driven feature selection enhances predictive performance.

Class Distribution Analysis

The distribution of the actual class labels (low, mid, and high risk) was compared to the expected distribution of the class obtained by the HyCIRRF model before assessing the model. Figure 6 demonstrates the real distribution of classes of the test data, and Figure 7 demonstrates the expected distribution.

The practically equal distribution of the real and projected numbers indicates that the hybrid ensemble preserves the balance of classes and does not have a bias in predicting the risk level of any specific level.

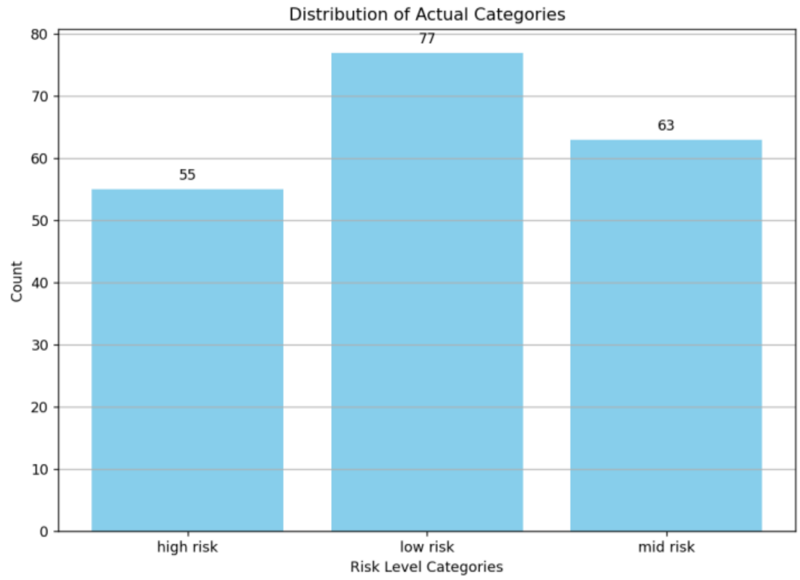

Figure 6. Actual Class Distribution

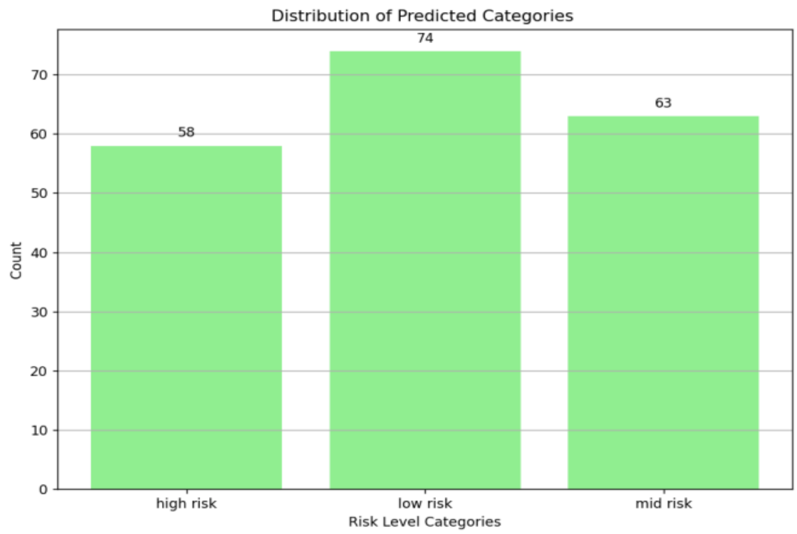

Figure 7. Predicted Class Distribution

Model Evaluation Metrics

The model developed was evaluated with five popular measures, such as accuracy, precision, and recall, F1-score, and error rate. The respective results are provided in Table 8, and the corresponding graphical analysis is provided in Figure 8.

Table 8. Performance Metrics of the Proposed HyCIRRF Model

Metric	Value (%)
Accuracy	93
Precision	93
Recall	94
F1-Score	93
Error Rate	7

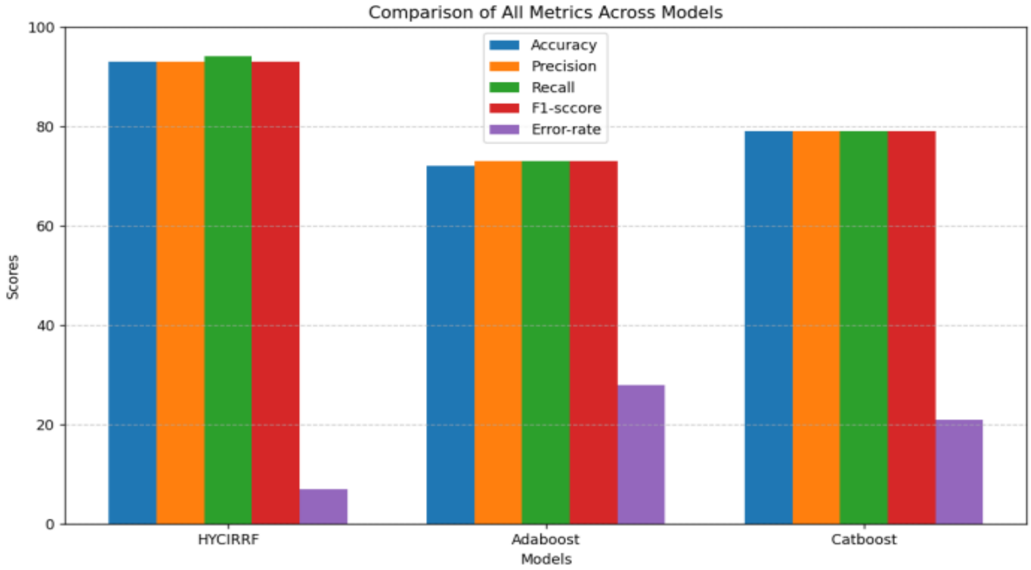

Figure 8. Comparison of all Model Performance Metrics between the proposed HyCIRRF model and the baseline models

The HyCIRRF model shows strong and balanced performance with an overall accuracy of 93% (Precision = 93%, Recall = 94%, F1 = 93%, Error Rate = 7%) (Figure 7). This reflects a balanced relationship between false positives and false negatives. By the tight alignment of accuracy and recall, which is important for clinical risk assessment?

Confusion Matrix Interpretation

Figure 9 gives a closer look at the findings of the classification of each of the three risk categories.

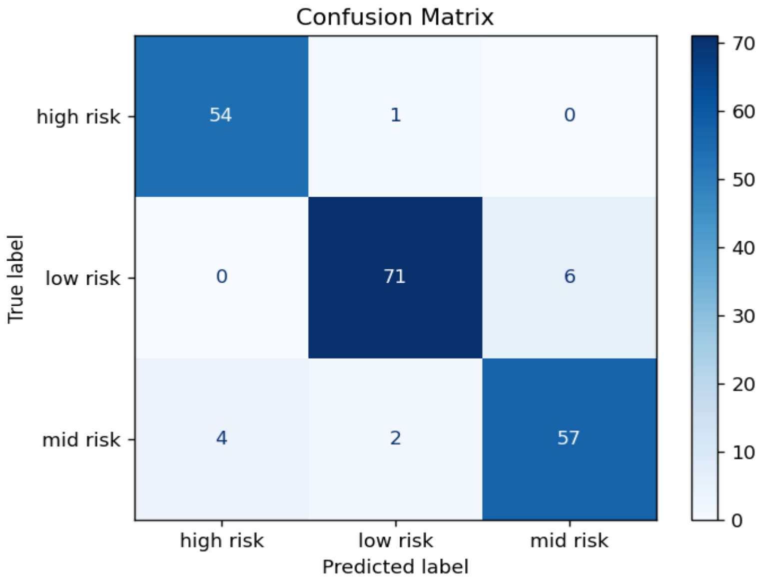

Figure 9. Confusion Matrix of the proposed model

Diagonal dominance implies that correct classification is likely in most situations, but low levels of off-diagonal inputs imply that there are few misclassifications between adjacent levels of risk. This finding indicates the strength of the model in recognizing the differences in the two types of pregnancy, namely, the medium-risk and the high-risk pregnancies, which are often similar in clinical aspects.

ROC Curve and AUC Evaluation

The ROC analysis was conducted with the help of a one-vs-rest approach to each risk category. Figure 10 illustrates the multi-class ROC curves, and in all classes, the corresponding Area Under Curve (AUC) values are above 0.95. These results prove that the HyCIRRF model is able to distinguish high-risk and non-high-risk pregnancies, which proves its outstanding discriminating ability.

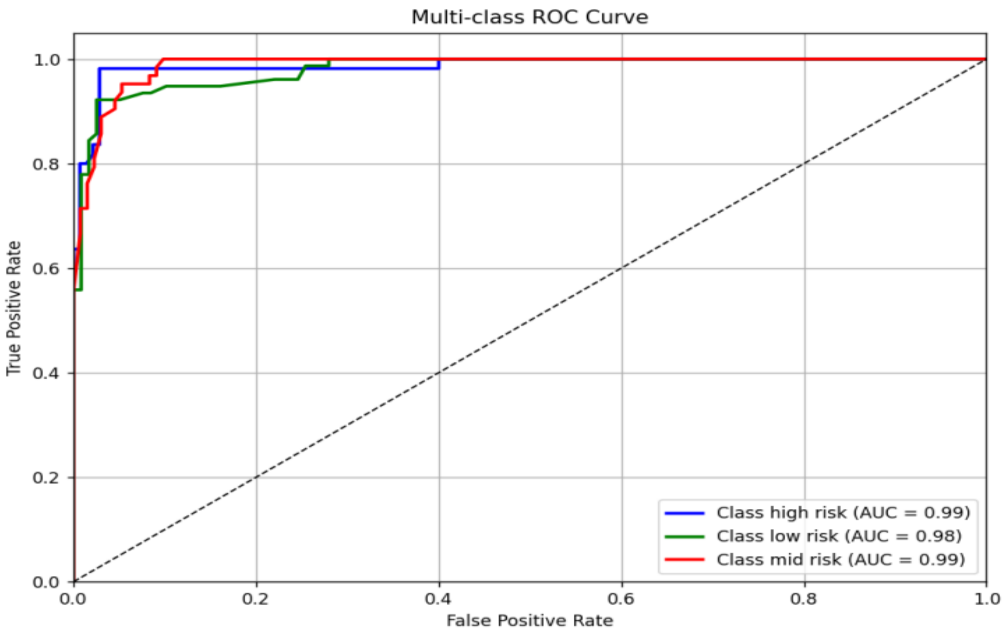

Figure 10. Multi-Class ROC Curve and AUC

Calibration Analysis

Calibration of probabilities is as important in clinical risk prediction as classification accuracy. The Brier Score of the proposed model is: 0.0378. A Brier score of less than 0.05 means that the quality of calibration is high. The calibration curve shows that there is a good agreement between the predicted probabilities and the observed results, which proves the high-quality probabilistic estimation.

This finding reinforces the clinical relevance of the HyCIRRF model in Figure 11 because well-calibrated models minimise the chances of an overconfident misclassification in medical decision support systems.

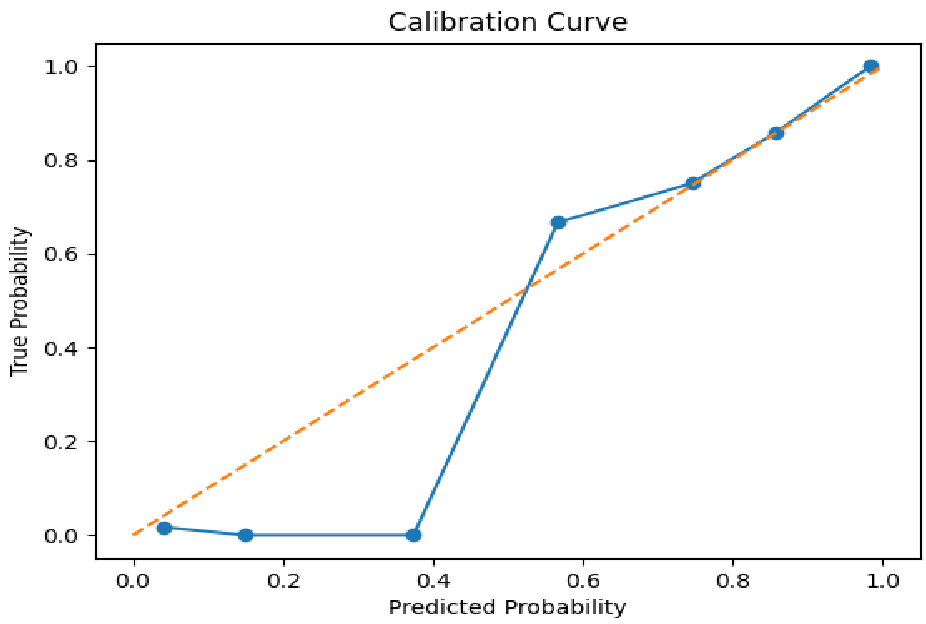

Figure 11. Calibration Analysis

Comparative Analysis with Baseline Models

Traditional machine learning classifiers were compared with the same evaluation platform, including Logistic Regression, Decision Tree, SVM, and XGBoost. Table 9 explains the results of the comparison. Strong diagonal dominance in the confusion matrix and AUC values greater than 0.95 for all risk categories demonstrate reliable multi-class risk stratification, thereby validating Hypothesis H4.

Table 9. Comparative Performance of Different Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Error Rate (%)
Logistic Regression [25]	78.0	65.0	1.00	79.0	22.0
Decision Tree [25]	67.0	56.0	91.0	69.0	33.0
SVM [26]	82.0	83.0	86.0	84.0	18.0
XGBoost [26]	76.0	77.0	82.0	79.9	24.0
DT+PSO	77.1	67.1	43.3	52.6	22.0
BiLTCN	89.0	83.0	82.0	82.0	11.0
SVM(RBF)+PSO	86.0	83.3	81.7	83.3	14.0
HyCIRRF (Proposed) Model	93.0	93.0	94.0	93.0	7.0

In every evaluation metric, the HyCIRRF model performed competitively with the baseline models, attaining 93 % accuracy and maintaining balanced precision and recall values. The ensemble soft-voting mechanism and CSO-based feature optimization contributed to stable predictions and reduced classification bias. The comparative analysis shows that the proposed HyCIRRF model has performance levels that are similar to strong baseline classifiers with a balance in precision and recall. This observation along with the repeated cross-validation analysis confirms Hypothesis H3 by establishing competitive and stable ensemble performance.

Ablation Study

To assess the role of each component in the HyCIRRF model, we performed an ablation study by omitting important components such as FastICA (ICA), Crow Search Optimization (CSO), and the ensemble learning approach as in Table 10.

- Model without ICA: When ICA was removed for extraction of features, the performance was 87% accuracy with an error rate of 0.13. This indicates that feature extraction does have a role but does not drastically affect the model.
- Model without CSO: With No CSO for feature selection the model still achieved 87% with an error rate of 0.13, indicating that feature selection is helpful, but not essential.
- Model without Ensemble: Removing ensemble learning (using only a single classifier) also resulted in 87% accuracy, with 0.13 error rate, indicating that ensemble learning is helpful but not a significant factor in this case.
- HyCIRRF (Proposed) Model: The full HyCIRRF model, with the use of ICA, CSO, and ensemble learning, achieved 93% accuracy, 94% recall, and 0.07 error rate, which proved the superior performance when combining all the components.

Table 10. Ablation study of proposed method

Model	Accuracy	Precision	Recall	F1 Score	Error Rate
Without ICA	0.87	0.87	0.87	0.87	0.13
Without CSO	0.87	0.87	0.87	0.87	0.13
Without Ensemble Learning	0.87	0.87	0.87	0.87	0.13
HyCIRRF (Proposed) Model	0.93	0.93	0.94	0.93	0.07

The ablation study highlights the importance of combining ICA, CSO and ensemble learning to get the highest predictive performance with the HyCIRRF model outperforming individual components by a large margin.

Statistical Robustness and Inferential Validation

Although the proposed HyCIRRF model demonstrated strong hold-out performance (93% accuracy), additional statistical validation was conducted to address robustness, reproducibility, and inferential significance concerns.

Repeated Stratified 10×5 Cross-Validation

Repeated Stratified 10-fold Cross-Validation with 5 repetitions was conducted to minimise bias related to single train-test splits. This test gives a more conservative and statistically sound estimate of generalisation performance. The repeated validation results were:

- Proposed HyCIRRF Accuracy: 0.8489 ± 0.0321

- Baseline Random Forest Accuracy: 0.8505 ± 0.0338

The proposed model has a 95 percent confidence interval of (0.8397, 0.8580).

The standard deviation is relatively small (about 3%), which means that the performance is consistent when using various randomised partitions. The small confidence interval also indicates statistical consistency in spite of the moderate size of the dataset. It should be mentioned that the 93% accuracy above is optimised hold-out evaluation, and repeated cross-validation gives a conservative estimate of robustness in a variety of resampling conditions.

Statistical Significance Testing

To determine whether performance differences between models were statistically meaningful, paired inferential tests were conducted across repeated folds.

The paired t-test and the Wilcoxon test in table 11 both reported p-values more than 0.05, which means that the difference in accuracy between the proposed HyCIRRF model and the baseline Random Forest model is not statistically significant.

Table 11. Statistical Significance Testing

Test	p-value
Paired t-test	0.5754
Wilcoxon Signed-Rank	0.4559
McNemar Test	1.000

Equally, the McNemar test (p = 1.000) implies that the two models have similar patterns of misclassification at the instance level.

These findings suggest that the suggested model has competitive and statistically stable performance instead of statistically superior performance in repeated cross-validation.

The cross-validation analysis has shown that the proposed HyCIRRF model has a competitive predictive performance (0.8489 ± 0.0321) in comparison with the baseline model (Random Forest) (0.8505 ± 0.0338). The paired t-test (p = 0.5754) and the Wilcoxon signed-rank test (p = 0.4559) show that there is no statistically significant difference between the models. This validates the hypothesis that the hybrid soft-voting ensemble is statistically stable and comparably performing on a series of randomised folds and thus Hypothesis H3 is supported in the sense of competitive and reliable predictive performance as opposed to statistical superiority.

Overall Statistical Interpretation

The additional robustness analyses confirm that:

- The HyCIRRF model maintains stable performance under repeated cross-validation.
- Statistical testing indicates that the proposed model demonstrates statistically comparable performance to the baseline model under repeated cross-validation, confirming robustness rather than statistical superiority.
- The model is well-calibrated and clinically reliable.

- Fast-ICA is justified due to severe multicollinearity in physiological predictors.

Together, these findings reinforce the validity of Hypotheses H1–H4 and strengthen the statistical credibility of the proposed framework.

DISCUSSION

HyCIRRF model has a high predictive accuracy of preterm pregnancy with an accuracy of 93, precision of 93, recall of 94 and error rate of 7. These findings suggest a strong model that has a balanced sensitivity and specificity, thus it is very useful in predicting preterm pregnancy. Moreover, the ROC-AUC of all the categories of pregnancy risks exceed 0.95, which demonstrates the excellent discriminative power of the model, especially in distinguishing between low-, mid-, and high-risk pregnancies.

The HyCIRRF model has a clear edge over a number of baseline models, such as the Logistic Regression (LR) (accuracy of 78, recall of 1.00) and the Decision Trees (DT) (accuracy of 67, precision of 56) models. The improvement in performance is credited to the Fast Independent Component Analysis (Fast-ICA) used to extract features, Crow Search Optimization (CSO) used to select features, and the hybrid combination of Regression Random Forest (RRF) and Logistic Regression (LR). This combination leads to more stable and generalizable predictions, overcoming the overfitting and computational inefficiencies which are common in SVM (82% accuracy) and XGBoost (76% accuracy).

One of the major strengths of the HyCIRRF model is its interpretability that is important in clinical practise. Fast-ICA is important in eliminating multicollinearity among clinical features, stabilising the performance of the model. Moreover, the analysis of feature importance provided by CatBoost shows that such variables as systolic blood pressure and blood glucose levels are very important and are closely related to pregnancy risk. CSO guarantees the best feature selection by striking a balance between predictive capability and computational efficiency, which is why the model is applicable to resource-constrained healthcare settings. The ensemble approach makes the model more stable such that HyCIRRF model is reliable in different situations.

The study however has a number of limitations. The model was trained and tested using one dataset of 1,014 records which might not be representative of a wide population or a healthcare environment. Also, the dataset lacks imaging data or biomarkers, which may also enhance the accuracy of prediction. The next step in working should be to validate the model with multi-centre or longitudinal data and include other biomarkers like C-reactive protein (CRP) or interleukin-6 (IL-6). Development of the model in terms of temporal variations in maternal health indicators would help the model develop dynamic capabilities, which would enhance its predictive accuracy in the long run.

SUMMARY AND CONCLUSION

The paper presents Hybrid Crow-ICA Regression Random Forest (HyCIRRF) model that predicts the risk of preterm pregnancy. It combines Fast Independent Component

Analysis (Fast-ICA) to reduce multicollinearity, CatBoost to rank features in terms of importance, Crow Search Optimization (CSO) to select the most optimal features, and a soft-voting ensemble of Random Forest and Logistic Regression to do final classification. The model was tested on the Maternal Health Risk dataset, and it had 93 percent accuracy, 94 percent recall, 93 percent precision, and 7 percent error rate with multi-class ROC-AUC values of over 0.95 in all the risk categories, indicating that the model has a high discriminative power.

In order to achieve robustness, cross-validation (10x5) was repeated, and the results demonstrated competitive performance (0.8489 +0.0321 accuracy) with a 95 percent confidence interval of (0.8397, 0.8580). The baseline Random Forest model also demonstrated the same performance (0.8505 ± 0.0338), and there was no statistically significant difference between them (paired t-test p = 0.5754, Wilcoxon p = 0.4559, McNemar p = 1.000), which means that it is stable and reliable.

The application of fast-ICA was justified by the fact that it minimised the multicollinearity (VIF > 80) between the predictors. The calibration analysis gave a Brier score of 0.0378, which is a confirmation that there is strong probabilistic estimation. The ablation experiment emphasised the fact that the entire HyCIRRF system, which is a combination of ICA, CSO, and ensemble learning, performed much better than the individual systems.

Even though these findings are founded on a single dataset, which indicate that the HyCIRRF framework has the potential to offer clinically interpretable decision support in preterm pregnancy risk. The external validation and SHAP-based interpretability analysis should be incorporated into future work to increase clinical applicability.

NOMENCLATURE

HyCIRRF	Hybrid Crow-ICA with Regression Random Forest
FastICA	Fast Independent Component Analysis
CSO	Crow Search Optimization
EL	Ensemble Learning
SVM	Support Vector Machine
LR	Logistic Regression
RF	Random Forest

AUTHOR CONTRITUTIONS

Conceptualization, P.G and S.A and P.G; methodology, P.G and S.A; software, P.G ;validation ,P.G and S.A ;formal analysis, P.G; investigation, P.G and S.A ;resources, P.G and S.A ;data curation, P.G ; writing- original draft preparation, P.G; writing-review and editing, P.G and S.A ;visualization, S.A ; supervision., P.G.

CONFLICT OF INTERESTS

The authors have no competing interests to declare that are relevant to the content of this research paper.

REFERENCES

1. World Health Organization. Preterm birth. Available from: https://www.who.int/news-room/fact-sheets/detail/preterm-birth, (Accessed on 11.11.2025).
2. Kassahun E. A. et al. Development and validation of a simplified risk prediction model for preterm birth: a prospective cohort study in rural Ethiopia, *Scientific Reports*, **2024**, *14*(1), 4845.
3. Lee, K.-S.; Ahn, K.H. Application of Artificial Intelligence in Early Diagnosis of Spontaneous Preterm Labor and Birth. *Diagnostics* **2020**, *10*, 733.
4. Kopanitsa, G., Metsker, O., Kovalchuk, S. Machine Learning Methods for Pregnancy and Childbirth Risk Management. *J. Pers. Med.* **2023**, *13*, 975.
5. Kloska A., Harmoza A., Kloska S. M., Marciniak T. and Sadowska-Krawczenko I. Predicting preterm birth using machine learning methods, *Scientific Reports*, **2025**, *15*(1), 5683.
6. Yu Q.-Y., Lin Y., Zhou Y.-R., Yang X.-J. and Hemelaar J. Predicting risk of preterm birth in singleton pregnancies using machine learning algorithms, *Frontiers in Big Data*, **2024**, *7*, 1291196.
7. Li J. Ge, J., Ran, N. et al. Finding the priority and cluster of inflammatory biomarkers for infectious preterm birth: A systematic review, *Journal of Inflammation*, **2023**, 20(1), 25.
8. Gupta K. and Alfirevic A. Systematic review of preterm birth multi-omic biomarker studies, *Expert Reviews in Molecular Medicine*, **2022**, 24, e18.
9. Williams S.M. et al. Defining knowledge gaps in preterm birth research: Can biomarkers fill the gaps?, *Frontiers in Medicine*, **2025**, 12, 1655833.
10. Włodarczyk, T., Płotka, S., Szczepański, T., Rokita, P., Sochacki-Wójcicka, N., et al. Machine learning methods for preterm birth prediction: a review. *Electronics*, **2021**, *10*(5), 586.
11. Sharifi-Heris, Z., Laitala, J., Airola, A., Rahmani, A. M., & Bender, M. Machine learning approach for preterm birth prediction using health records: systematic review. *JMIR Medical Informatics*, **2022**, *10*(4), e33875.
12. Raja, R., Mukherjee, I., & Sarkar, B.K. A machine learning-based prediction model for preterm birth in rural India. *Journal of Healthcare Engineering*, **2021**, *2021*(1), 6665573.
13. Desuky, A. S., Hussain, S., Cifci, M. A., El Bakrawy, L. M., Mzoughi, O., & Kraiem, N. (2024). Parameter Optimization Based Mud Ring Algorithm for Improving the Maternal Health Risk Prediction. *IEEE Access*, **2024**, 12, 167245- 167261.
14. Yalçın, E., Koç, T.K., Aslan, S., Demir, S.C., Aykut, S., & Sucu, M. Hyperparameter optimization based machine learning approach for early diagnosis of fetal genetic disorders. *Discover Computing*, **2025**, *28*(1), 274.
15. Moreira, M.W., Rodrigues, J.J., Kumar, N., Al-Muhtadi, J., & Korotaev, V. Nature-inspired algorithm for training multilayer perceptron networks in e-health environments for high-risk pregnancy care. *Journal of Medical Systems*, **2018**, *42*(3), 51.

16. Pavanya, M., Chadaga, K., Vasudeva, A., Rao, B.K., Prabhu, S., & Bhat, S.K. Prediction of birthweight with early and mid-pregnancy antenatal markers utilising machine learning and explainable artificial intelligence. *Scientific Reports*, **2025**, *15*(1), 26223.
17. Du, Y., Rafferty, A.R., McAuliffe, F.M., Wei, L., & Mooney, C. An explainable machine learning-based clinical decision support system for prediction of gestational diabetes mellitus. *Scientific Reports*, **2022**, *12*(1), 1170.
18. Khan, W., Zaki, N., Ghenimi, N., Ahmad, A., Bian, J., Masud, M. M., at al. Predicting preterm birth using explainable machine learning in a prospective cohort of nulliparous and multiparous pregnant women. *PLoS One*, **2023**, *18*(12), e0293925.
19. Zhang F., Tong L., Shi C., Zuo R., Wang L., and Wang Y. Deep learning in predicting preterm birth: A comparative study of machine learning algorithms, *Maternal-Fetal Medicine*, **2024**, *6*(3), 141-146.
20. Pradipta G.A., Wardoyo R., Musdholifah A. and Sanjaya I. N. H. Improving classification performance of fetal umbilical cord using combination of SMOTE method and multiclassifier voting in imbalanced data and small dataset, *Int. J. Intell. Eng. Syst.*, **2020**, *13*(5), 441-454.
21. Qi. S., Xiaoxuan, Z., Yousheng, Y., et al. Machine learning-based prediction model of preterm birth using electronic health record, *Journal of Healthcare Engineering*, **2022**, *2022*, 1-12.
22. Siargkas, A., Tsakiridis, I., Kappou, D., Mamopoulos, A., Papastefanou, I., Dagklis, T. Machine Learning Models for the Prediction of Preterm Birth at Mid-Gestation Using Individual Characteristics and Biophysical Markers: A Cohort Study. *Children* **2025**, *12*, 1451.
23. Ayu, P.D.W., Pradipta, G.A., Susila, I.M.D., Hostiadi, D.P. and Liandana, M. Deep learning-based detection and classification of amniotic fluid echogenicity type for enhanced prenatal diagnosis, *Int. J. Intell. Eng. Syst.*, **2025**, *18*(1), 246-267.
24. Joloudari, J.H., Marefat, A., Nematollahi, M.A., Oyelere, S.S., Hussain, S. Effective Class-Imbalance Learning Based on SMOTE and Convolutional Neural Networks. *Appl. Sci.* **2023**, *13*, 4006.
25. Saroj, R.K., & Anand, M. Environmental factors prediction in preterm birth using comparison between logistic regression and decision tree methods: An exploratory analysis. *Social Sciences & Humanities Open*, **2021**, *4*(1), 100216.
26. Tomar, K., Sharma, C.M., Prasad, T., Chariar, V.M. A Machine Learning-Based Risk Prediction Model During Pregnancy in Low-Resource Settings. *Med. Sci. Forum* **2024**, *25*, 13.