Svedberg Open
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Explainable AI For Predictive Modelling of Maternal Health Behavior Under PMMVY: Accessibility, Benefit Adequacy, and District Heterogeneity

Dr. Happy Agrawal[1], Dr Shruti Rawal[2], Dr. Megha Choudhary[3]

[1]Assistant Professor, Department of Business Administration, St. Xavier's College Jaipur, Email ID: guptahappy29@gmail.com ORCID: 0000-0001-7247-668X

[2]Assistant Professor, Department of English, St. Xavier's College Jaipur, Email ID: shrutivipul@gmail.com, ORCID: 0000-0002-2354-5631

[3]Lecturer, Department of English, Government Polytechnic College, Bhilwara, Email ID: meghachoudharyy@gmail.com

ABSTRACT

This research proposes an explainable predictive model to evaluate maternal health behavior of beneficiaries of Pradhan Mantri Matru Vandana Yojana (PMMVY) in the districts of Karauli and Dholpur of Rajasthan, India. The analytical sample was cross-sectional of 350 beneficiaries in rural areas, equally distributed between the two districts. Maternal health behaviour was assessed by a three item composite score, and predictor variables were the demographic characteristics, four indicators of PMMVY accessibility, and perceived adequacy of programme benefits. Conventional benchmark models were evaluated using linear regression, Ridge regression, Elastic Net and a dummy regressor and nonlinear predictive models were assessed with repeated nested cross validation. The final Gradient Boosting model had an MAE of 0.208, an RMSE of 0.535 and an R^2 of 0.865, which is better than the best conventional model. Digitalisation related ease of enrolment was the most important predictor according to Permutation Feature Importance and SHAP analysis, followed by education and ease of documentation. The additional predictive value of perceived benefit adequacy was small. Pooled performance was relatively stable across predictor-profile grouping in robustness analyses, but there was considerable district heterogeneity and poor cross-district transportability. Generally, accessibility factors were more predictive than perceived benefit adequacy. The results highlight the importance of explainable predictive analysis for programme level maternal health assessment and the need for external validation prior to wider geographic application.

Keywords: Pradhan Mantri Matru Vandana Yojana (PMMVY); Explainable Predictive Modelling; Maternal Health Behavior; Healthcare Accessibility; Gradient Boosting.

INTRODUCTION

Access to health services is not sufficient to promote maternal health behaviour. Women should also have access to, afford and be able to use maternal health services during the ante- and post-natal period. Despite the rollout of maternal health programmes and financial-support interventions, significant inequities in the use of antenatal, delivery, and postnatal services exist between socioeconomic and geographic groups in India (Mehta et al., 2023; Chi et al., 2024). The national evidence also shows that the continuity of maternal health services is not consistent and is affected by household characteristics, socioeconomic conditions, and the environment of the service (Longchar et al., 2025). The patterns highlight the need to consider access and financial assistance as two separate components when evaluating maternal health behaviour.

Pradhan Mantri Matru Vandana Yojana (PMMVY) is a national maternity benefit scheme to support women during pregnancy and early motherhood. Current findings indicate that cash transfers for mothers can have positive impacts on the use of maternal and child health services, but their impact is context and service delivery dependent. Banerjee et al. (2025) found that the participation in PMMVY was linked to better utilization of maternal and child health services, and Jagannath and Chakravarthy (2025) reported that PMMVY was related to better access to services and health outcomes for mothers and children. The relevance of PMMVY to the policy is also evident from the evidence which shows that the programme has been associated with child anthropometric outcomes (Ray et al., 2026). Furthermore, cash-plus interventions suggest that financial support could work better if supported by behavioural and service support mechanisms (Marchant et al., 2026; Mishra & Vellakkal, 2026).

But the receipt of programme benefits does not mean that the health care access is effective or financial support for pregnancy-related needs is adequate. The relevance of maternity benefits to the healthcare cost may depend on whether the transfer is helpful in offsetting healthcare cost (Mathur & Sen, 2025). The situation is particularly important in Rajasthan where women are more likely to face physical, financial, administrative and social constraints to accessing maternal health services (Casebolt et al., 2023). Thus, a programme enrolment or benefit receipt-based assessment could miss out on important differences in accessibility and perceived benefit adequacy.

Another challenge is the types of analysis typically employed to analyse maternal health care utilization. While conventional linear models are useful for estimating average associations, they may not be sufficient to capture the nonlinearity, threshold effects and interactions between demographic characteristics, service accessibility, administrative procedures, digital enrolment, frontline-worker assistance and perceived benefit adequacy. Advanced predictive modelling offers a valuable tool for analyzing such complicated relationships and highlights out-of-sample predictive performance. However, the predictive accuracy is not enough for policy interpretation. Analytical methods that are explainable must be used to determine which predictors are most important to model performance and if there are predictive patterns that are consistent across geographic contexts.

Thus, the present study uses an explainable supervised predictive approach to evaluate maternal health behaviour among the beneficiaries of PMMVY in the districts of Karauli and Dholpur in Rajasthan. Accessibility is reflected in physical access to service points, assistance provided by frontline workers, ease of documentation and ease of digitalisation-related enrolment, with perceived benefit adequacy assessed as an additional programme-related predictor. The study is predictive, not causal, and primarily examines model performance, model interpretability, model robustness, and geographic transportability.

The study has three objectives:

- To develop and compare predictive models for maternal health behaviour using demographic, accessibility, and benefit-adequacy variables.
- To assess the predictive importance of PMMVY accessibility and perceived benefit adequacy using permutation feature importance, SHAP, and ablation analysis.
- To evaluate model robustness and geographic transportability across Karauli and Dholpur.

2. Literature Review

Maternal health behaviour refers to the practice of registering during pregnancy, attending recommended check-ups, delivering at a health facility, postnatal care, and continuity of services. Despite the growth of public maternal health programmes in India, utilization is uneven across socioeconomic status, education, residence, parity, and health-system characteristics (Girotra et al., 2023). A lack of early initiation of antenatal care may limit early detection of risk and reduce ongoing participation in formal health care. The availability of services does not mean that they are effectively used as there is a significant difference in the timing of the first antenatal visit as confirmed by the national evidence (Tripathy & Mishra, 2023).

Financial incentives are intended to reduce demand-side barriers. The effects of conditional maternity assistance on household investments in maternal and child health vary across social and demographic groups (Kekre & Mahajan, 2023). Partly, its behaviour effect is determined by the assessment of the benefit value in comparison to the actual costs. Even though institutional delivery has seen a significant reduction in out-of-pocket expenses, the costs still differ from public to private institutions (Manna et al., 2023). The financial burden of a c-section could also be an added strain, meaning that maternity benefits might not include transportation, medicines, diagnostics, nutrition, and other pregnancy-related expenses (Singh et al., 2023). So, the perception of benefit adequacy is not the same as receipt of benefits.

Another key factor in maternal health care behaviour is accessibility. Coverage improvements have not yet led to a reduction in geographic inequalities in maternal-care opportunities (Dandona et al., 2024). Household characteristics and local service conditions are also factors that influence rural utilisation (Tripathi et al., 2024). Results from more than 40,000 mother-child dyads demonstrate ongoing losses in maternal, newborn and child health care (Ghosal et al., 2025). Dropout from pregnancy to postpartum care is also linked to socioeconomic disadvantage (Singh et al., 2025). Accessibility is not just about distance, but also about administrative procedures, documentation, frontline assistance and continuity of health-system contact. Digital supportive supervision could help address some barriers and enhance utilisation (Sharda et al., 2026).

In recent years, machine learning has been used in the prediction of maternal health care. Researchers have predicted the best antenatal-care utilisation, healthcare access and key determinants in rural areas (Sani et al., 2025; Kassaw et al., 2025). Hierarchical models have also identified determinants that act at various levels of analysis (Osborne et al., 2026). These methods can detect nonlinearities, interactions and heterogeneous patterns of predictor variables that are not possible to detect with traditional methods.

Interpretability of models has become more relevant. Machine learning and SHAP analysis have been employed to predict maternal-care completion, non-utilisation of postnatal-care, and antenatal-care dropout (Noor et al., 2026; Dhar et al., 2026; Kebede et al., 2025; Yoseph et al., 2026). This is a change from accuracy in prediction to transparent and policy-relevant modelling.

But there is a lack of research that combines perceived adequacy of maternity benefits and programme accessibility in one explainable framework for PMMVY beneficiaries. There is also limited evidence of the stability of the predictors, incremental programme-related value and cross-district transportability. The current research aims to fill these gaps by integrating comparative predictive modelling, permutation importance, SHAP interpretation, ablation analysis, robustness assessment and cross-district validation in both the districts of Karauli and Dholpur. It therefore not only assesses predictive performance, but also interpretability, stability and geographic generalisability.

3. Methodology

3.1 Research Design

The present study used a cross sectional supervised predictive-modelling design to predict Maternal Health Behavior (MHB) among beneficiaries of Pradhan Mantri Matru Vandana Yojana (PMMVY) of Karauli and Dholpur districts of Rajasthan, India. A total of 350 rural beneficiaries were selected, 175 beneficiaries per district. Pooled predictive modelling and district-specific robustness and transportability analyses were possible due to the equal representation of the two districts. This balanced analytical composition, however, was not taken as a sign of representativeness of the population.

The study was a predictive one and not causal. It was mainly designed to arrive at an estimate of MHB based on demographic characteristics, dimensions of PMMVY accessibility and perceived adequacy of programme benefits. As a result, out-of-sample prediction, comparative model performance, model interpretability and robustness were highlighted in model development. The associations and feature-importance estimates were drawn in a predictive sense only, and no causal inferences were drawn. There were more records in the source workbook than were included in the final analytical sample that had been labelled by the district. Additional observations were made, but the available analytical data did not contain enough information to confirm the sampling or exclusion criteria for these observations. Thus, no assumption was made about their exclusion, and all analyses were limited to the pre-determined sample of 350 beneficiaries.

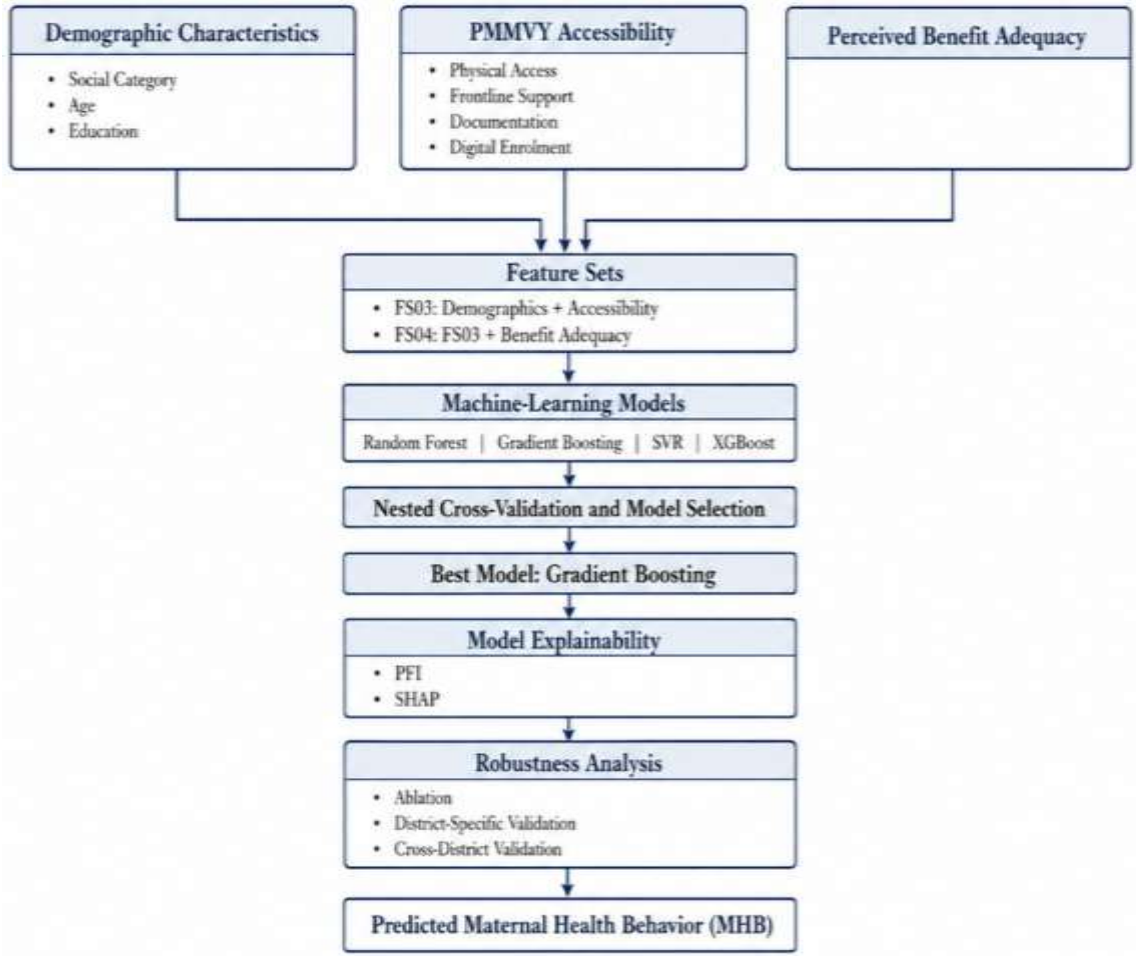

Figure 1. Simplified analytical framework for explainable machine-learning prediction of maternal

health behavior among PMMVY beneficiaries.

3.2 Description of data set, variables and pre-processing

The dependent variable was Maternal Health Behavior, which was measured by a composite score from three five-point likert scale items related to maternal and maternal-child health practices. The items evaluated diet and nutrition, child vaccination and routine health checks, and self-efficacy in maternal and child health management. Three item scores were added together to give the resulting mhb_score, which ranges from 3 to 15. The observed values were between 11 and 15. The construct showed good internal consistency (Cronbach's $\alpha = 0.945$). A large ceiling effect was noted, with over 50% of the respondents scoring at the highest level. This limited outcome variability was taken into account when interpreting predictive performance, including district-specific R^2 values.

Four five-point indicators were used to represent PMMVY accessibility: accessibility to an Anganwadi centre or local government service point, assistance from ASHA or Anganwadi workers, ease of completing documentation, and ease of enrolment due to digitalisation. The four indicators showed good internal consistency (Cronbach's $\alpha = 0.963$). The composite accessibility score was only used descriptively. The individual accessibility indicators were kept separately for predictive modelling so that feature level interpretability was maintained.

Perceived benefit adequacy was measured with a single ordinal item that asked whether the financial assistance provided by PMMVY was adequate to address healthcare-related needs. Other factors were social category, age group, and level of education. Social category was coded as a nominal variable, while age group and education level were coded as ordered variables. District was deliberately omitted from the main predictive specification as the focus was to estimate MHB from attributes related to the beneficiary and the programme. District was then added to sensitivity and geographic transportability analyses.

To reduce leakage of information in the modelling process, pre-processing was applied. Only the training partitions were used for estimating all data transformations. Nominal categorical variables were one-hot encoded and ordinal variables had their order preserved. For categorical observations, missing data were imputed using the mode, and for numeric or ordinal observations, the missing data were imputed using the median. Only scale-sensitive models such as Support Vector Regression (SVR) and linear models were standardized. Variables that directly overlapped with MHB or were derived from the outcome were excluded. Predictive specifications were also not included for the composite accessibility score when the individual component indicators were included.

3.3 Baseline Models

The incremental predictive value of the nonlinear predictive approaches was compared to a set of conventional benchmark models. The baseline models included a dummy mean regressor, Ordinary Least Squares regression, Ridge regression and Elastic Net regression. These were all together the naive prediction, unregularised linear prediction, and regularised linear prediction.

The baseline models were pre-processed using the same pre-processing principles and outer validation partitions as were applied to the proposed predictive models. Only models that needed hyperparameter tuning were optimised using nested inner-loop optimisation. This approach allowed for the assessment of comparative performance based on observations that were not used in the optimisation of the hyperparameters, and it also did not require the tuning of models that did not have comparable hyperparameter requirements. The baseline models were used only as predictive models and not as causal explanatory models.

3.4 Proposed Predictive Models

To assess the extra predictive information provided by PMMVY accessibility and perceived benefit adequacy, a staged feature-set approach was used. Demographic-only specifications were considered in the initial modelling phase, and two main sets of features were then considered for the final comparative analysis. FS03 included the four individual PMMVY accessibility indicators with demographic characteristics, while FS04 added perceived benefit adequacy to FS03.

The four nonlinear predictive algorithms were tested: Random Forest, Gradient Boosting Regression, Support Vector Regression, and XGBoost. These algorithms have been chosen as they are of different functional structures and can model nonlinear relationships, threshold effects and interactions which may not be adequately captured by the traditional linear specifications.

The full predictive specification included social category, age group, education level, the four PMMVY accessibility indicators, and perceived benefit adequacy as candidate predictors of the target variable, MHB. The prespecified model-selection procedure was used to select the predictive algorithm. There was no a priori winning algorithm; the final model was chosen only by comparing the performance of different models using cross-validation and then reported in the Results section.

3.5 Model Evaluation, Explainability, and Robustness

To get robust estimates of predictive performance repeated nested cross validation was used. The outer validation procedure was five-fold cross-validation repeated 10 times, which resulted in 50 outer-fold evaluations. Hyperparameter optimisation was done using three-fold inner cross validation within each outer training partition. Using randomized search, 15 hyperparameter configurations were tested for each applicable predictive model. For computational reproducibility, a fixed random seed (42) was used.

The main model-selection criterion was the Root Mean Squared Error (RMSE). Complementary measures of predictive accuracy and stability were Mean Absolute Error (MAE), coefficient of determination (R^2), and variability in RMSE across the validation folds. The lower the RMSE and MAE, the less the prediction error, while the higher the R^2 the better the predictive fit. Since the MHB outcome had a high ceiling concentration, the MAE and RMSE were used in conjunction with the R^2, not just the R^2.

Permutation Feature Importance (PFI) and SHapley Additive exPlanations (SHAP) were used to evaluate model explainability. PFI was used to quantify the loss of predictive performance from random permutation of individual predictors and was repeated 20 times to provide robust estimates of predictor importance. TreeSHAP was employed to estimate the relative importance of individual predictors in a fitted prediction in the final tree-based predictive model. Therefore, PFI was used to measure the performance-based predictive importance and SHAP to measure the contribution of the features to the fitted predictive function. In neither case was there any indication of causal effects.

The benefit-adequacy ablation grouped cross-validation based on repeated predictor profiles, within-district evaluation, district-inclusion sensitivity analysis, and cross-district transfer testing were included in the robustness and sensitivity analyses. These analyses tested the hypothesis that repeated covariate structures, geographic composition, or district-specific distributions disproportionately affected the pooled predictive performance. Linear statistical benchmarks were also analyzed with HC3 heteroskedasticity-robust standard errors, as appropriate.

The final manuscript should include the exact computational environment, software and package versions, hyperparameter search spaces, and parameter settings used to obtain the results reported, directly from the analytical notebook used for the analysis. These details should not be reconstructed or assumed if the original analytical record is not available.

3.6 Ethical Considerations

Analysis was done on an existing anonymous dataset of PMMVY beneficiaries without any intervention or modification of healthcare services. Only variables directly relevant to the research objectives were analysed and no directly identifying information was included in the analytical dataset.

The analytical files available do not contain enough information to confirm the name of the ethics committee that approved the study, the ethical approval number, the original informed consent procedure, or the confidentiality protection measures at the field level. Thus, these information should be taken from the original study or project documentation and included in the final manuscript prior to submission to the journal. If there is no supporting information, then do not assume or add any ethical approval or consent information.

4. Results and Discussion

4.1 Descriptive and Construct Characteristics

The final analytical sample comprised 350 PMMVY beneficiaries, including 175 participants from Karauli and 175 from Dholpur. Table 1 presents the descriptive and measurement characteristics of the two principal multi-item constructs.

Table 1. Descriptive and Measurement Properties of the Principal Constructs

Construct	N	Items	Mean	SD	Observed Range	Cronbach's α	Ceiling, n (%)
PMMVY Accessibility	350	4	16.760	3.236	12–20	0.963	164 (46.86)
Maternal Health Behavior	350	3	13.737	1.460	11–15	0.945	185 (52.86)

The internal consistency of both constructs was high. The mean PMMVY accessibility score was 16.76 out of 20 while the mean MHB score was 13.74 out of 15. Both of these measures, however, had high ceiling concentrations. In particular, 52.86% of beneficiaries obtained the maximum MHB score. This limited outcome variation can have an influence on the error measures and R^2 and is important when interpreting the predictive performance, especially in district based analyses. Table 1A shows descriptive patterns at the district level. These comparisons are of special interest because district heterogeneity is a key part of the study's predictive framework.

Table 1A. District-Specific Distribution of Principal Study Variables

Variable	Karauli (n = 175)	Dholpur (n = 175)
Mean MHB score	12.48	14.99
Mean accessibility score	13.67 / 20	19.85 / 20
Mean perceived benefit adequacy	3.00 / 5	4.34 / 5
Maximum MHB score, n (%)	11 (6.29)	174 (99.43)
Maximum accessibility score, n (%)	0 (0.00)	164 (93.71)
Benefit adequacy distribution	All respondents scored 3	25 scored 1; 3 scored 2; 6 scored 4; 141 scored 5

There were district-level differences. The ceiling concentration of Dholpur was nearly 100% in MHB and accessibility while the ceiling concentration of Karauli was almost double the former. Perceived benefit adequacy was also found to be structurally different among the districts. While all the respondents from the village of Karauli selected the middle category, 141 respondents in Dholpur selected the maximum. These distributions include significant contextual variation and give strong arguments for the following within-district, district-sensitivity and cross-district transportability analyses.

4.2 Selection of the Proposed Predictive Model

Repeated nested cross-validation was used to compare the candidate predictive algorithms. FS03 categorized the four individual PMMVY accessibility indicators with the demographic characteristics, while FS04 further examined perceived adequacy of PMMVY benefits. Table 2 shows the predictive performance of the eight possible model-feature-set combinations.

Table 2. Nested Cross-Validated Performance of Candidate Predictive Models

Rank	Feature Set	Model	MAE	RMSE	R^2	RMSE SD
1	FS04	Gradient Boosting	0.208	0.535	0.865	0.018
2	FS03	Gradient Boosting	0.207	0.540	0.863	0.021
3	FS03	XGBoost	0.238	0.547	0.859	0.018
4	FS04	XGBoost	0.237	0.547	0.859	0.014
5	FS04	SVR	0.222	0.553	0.856	0.028
6	FS04	Random Forest	0.236	0.554	0.856	0.014
7	FS03	Random Forest	0.237	0.557	0.854	0.016
8	FS03	SVR	0.229	0.567	0.848	0.032

Note: RMSE was the prespecified primary criterion for model selection. MAE, R^2, and RMSE variability were used as complementary indicators of predictive performance and stability.

Gradient Boosting with FS04 had the lowest nested cross validated RMSE of 0.535 and the highest R2 of 0.865 and thus was chosen as the proposed predictive model. The MAE of 0.207 with Gradient Boosting and FS03 was slightly higher than the 0.208 with FS04, with the latter outperforming according to the prespecified primary criterion and having a slightly higher R^2. The small difference in the two versions of Gradient Boosting suggests that the perceived benefit adequacy added only limited additional predictive information over and above demographic and accessibility characteristics.

It is also found that the most successful results were given by nonlinear ensemble methods for the comparative results. The relatively small differences in the performance of several candidate models, however, suggest that the choice of model should be understood as one of models that are validated as predictors and not as the best algorithms in all possible scenarios.

4.3 Baseline versus Proposed Model Performance

The proposed FS04 Gradient Boosting model was compared to the best baseline model of Ridge regression with FS03 to check whether this model had added value to the conventional models. The held-out validation results are shown in Table 3.

Table 3. Performance Comparison of the Best Baseline and Proposed Model

Metric	Best Baseline: FS03 +	Proposed: FS04 + Gradient	Improvement

	Ridge	Boosting	
MAE	0.395	0.208	47.38% reduction
RMSE	0.637	0.535	16.02% reduction
R^2	0.809	0.865	+0.056
Model form	Regularised linear	Nonlinear ensemble	—
Benefit adequacy included	No	Yes	—

Note: Values represent predictive performance estimated from the corresponding held-out validation procedures. R^2 represents predictive fit and should not be interpreted as evidence of causal explanation.

The model proposed reduced the MAE by 47.38% and the RMSE by 16.02% from the best conventional model available. The predictive R2 went from 0.809 to 0.865, which is an absolute increase of 0.056. The results suggest that the nonlinear Gradient Boosting specification learned predictive information that was not as well described by the linear benchmark with regularisation.

The improvement also is relevant to the proposed explainable predictive framework; it indicates that the chosen model may be useful for prediction, while at the same time allowing interpretation at the feature level by using complementary explainability procedures.

4.4 Explainability of the Proposed Predictive Model

Permutation Feature Importance (PFI) and SHapley Additive exPlanations (SHAP) were used to interpret the final Gradient Boosting model. Both methods enabled the evaluation of the relevance of the predictors from complementary viewpoints. Table 4 provides worldwide ranking of the predictors.

Table 4. Global Predictor Importance in the Proposed Gradient Boosting Model

Predictor	Permutation Importance	PFI Rank	Mean |SHAP|	SHAP Rank
Digitalisation made enrolment easier	0.482	1	0.491	1
Education	0.257	2	0.217	3
Ease of completing documentation	0.120	3	0.265	2
ASHA/Anganwadi assistance	0.074	4	0.211	4
Age group	0.064	5	0.095	7
Ease of reaching service point	0.057	6	0.153	6
Perceived adequacy of PMMVY benefit	0.017	7	0.162	5
Social category	0.001	8	0.019	8

The ease of enrolment due to digitalisation was a major factor in both the explainability approaches, with a permutation importance of 0.482 and a mean absolute SHAP value of 0.491. This uniformity suggests that digital accessibility of enrolment had a significant impact on the overall predictive performance and on fitted predictions per se.

Ease of document completion and education was also influential, but the relative influence of these two factors differed between the two approaches. Documentation ease ranked third according to PFI but second according to SHAP, whereas education ranked second and third, respectively. In both the methods, the ASHA/Anganwadi assistance was found to be relevant for making predictions, ranking 4th in both the viewpoints of explainability.

The perceived benefit-adequacy showed a more intricate pattern. It was the fifth most important according to the mean absolute SHAP and seventh most important according to permutation importance. This suggests that benefit adequacy was a factor in some individual predictions, but caused relatively little degradation in overall model performance when information was disrupted. This means its incremental predictive contribution is more directly assessed via the ablation analysis in Section 4.5.

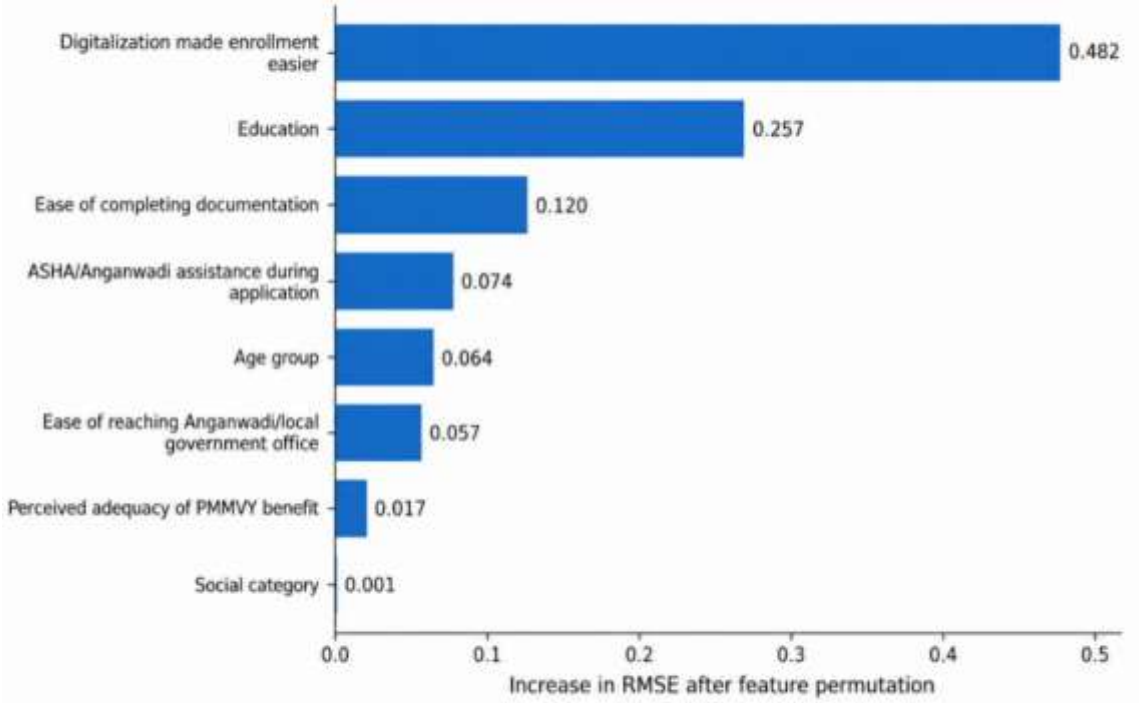

Figure 2. Permutation Feature Importance of Predictors in the Proposed Gradient Boosting Model.

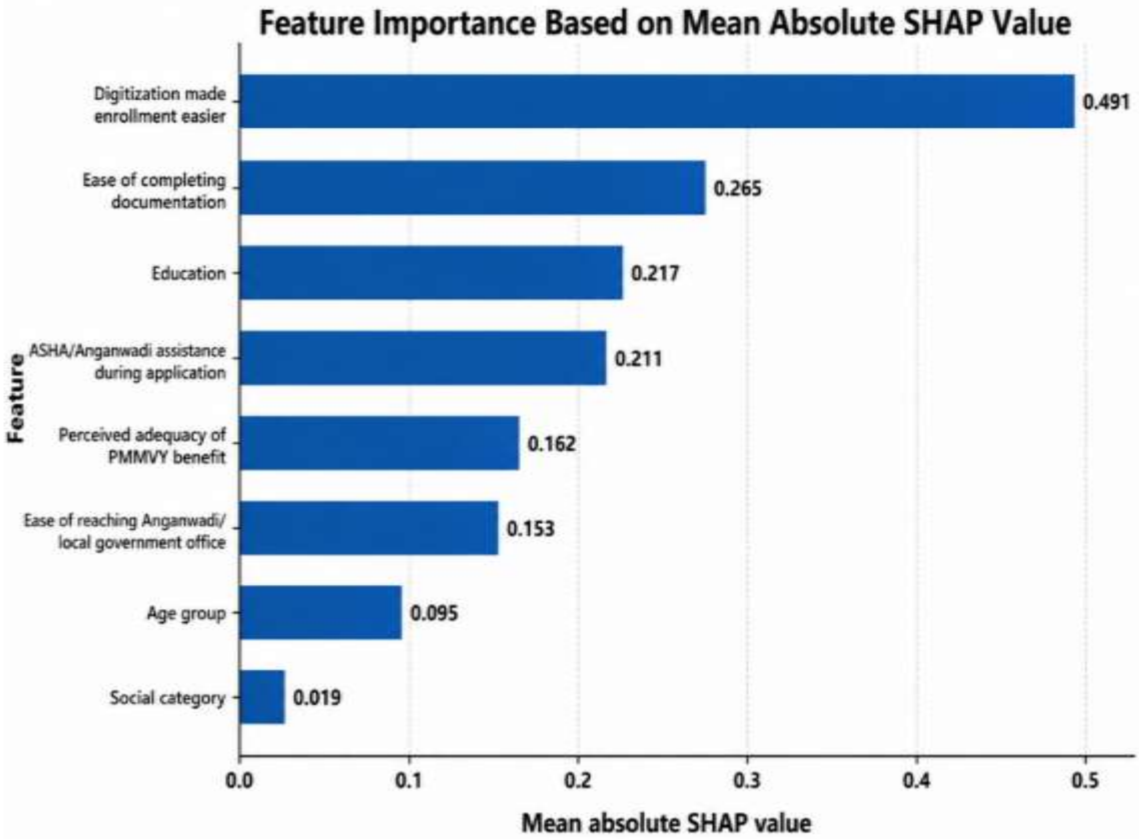

Figure 3. SHAP Global Feature Importance for the Proposed Gradient Boosting Model.

The overall results show that PFI and SHAP agreed on digitalisation-related enrolment accessibility as the main predictive factor. The rankings of the various secondary predictors illustrate the importance of the use of complementary explainability methods, as it is possible to find different dimensions of the relevance of a predictor.

4.5 Robustness, Sensitivity, and Geographic Transportability

Robustness analyses examined the stability of pooled predictive performance, the influence of repeated predictor profiles, district-specific performance, district sensitivity, cross-district transportability, and the incremental contribution of perceived benefit adequacy. Table 5 summarises the results.

Table 5. Robustness and Sensitivity Results for the Proposed Model

Analysis	Scenario	MAE	RMSE	R^2	Key Result
Primary nested CV	FS04 + Gradient Boosting	0.208	0.535	0.865	Primary estimate
Grouped nested CV	Predictor-profile grouping	0.245	0.550	0.857	Close to primary performance
Within-district	Karauli	0.345	0.700	0.547	Moderate predictive performance
Within-district	Dholpur	0.006	0.076	−0.006	Restricted outcome variation
Cross-district	Karauli → Dholpur	2.156	2.249	−889.267	Poor transportability
Cross-district	Dholpur → Karauli	2.520	2.727	−5.859	Poor transportability
District ablation	Without district	0.201	0.534	0.8659	Reference
District ablation	With district	0.189	0.512	0.8766	ΔR^2 = +0.0107
Adequacy ablation	Without adequacy	0.203	0.535	0.8653	Reference
Adequacy ablation	With adequacy	0.201	0.534	0.8659	ΔR^2 = +0.0006

The RMSE and R^2 in the grouped nested cross-validation were 0.550 and 0.857, respectively, compared with 0.535 and 0.865 in the primary pooled analysis. The relatively small deterioration suggests that the strong pooled predictive performance was not primarily attributable to repeated predictor profiles.

Predictive conditions were found to be quite different between districts. The model in Karauli had an RMSE of 0.700 and an R^2 of 0.547, which is considered moderate. In Dholpur, the RMSE was only 0.076, but R^2 was −0.006. In Dholpur, the apparently low prediction error should not be interpreted as better model performance since almost all beneficiaries achieved the maximum MHB score. The very narrow range of outcomes significantly reduced the possibility of discrimination between individuals.

Geographic distribution shift was supported by cross-district validation. The model was trained using the data from Karauli and tested with the data from Dholpur, which resulted in an RMSE of 2.249 and an R2 of −889.267. On the other hand, training on Dholpur and testing on Karauli gave RMSE value as 2.727 and R2 value as −5.859. The R^2 values for these relationships are all negative and strong, suggesting that relationships estimated in one district were not successfully transferred to the other. Thus, good internal predictive performance in a pool does not mean geographic generalisability.

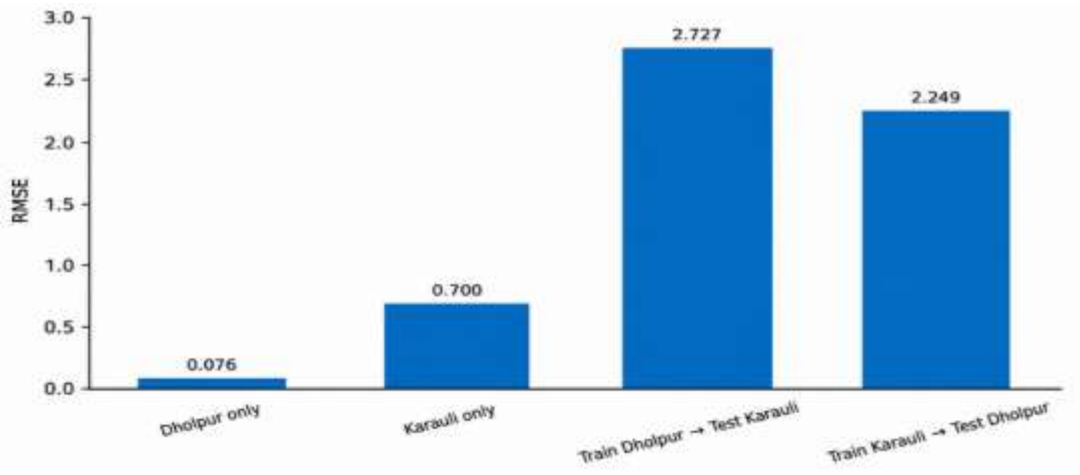

Figure 4. Geographic Robustness of the Proposed Model Based on RMSE.

Contextual heterogeneity was also confirmed by sensitivity analysis carried out at the district level. Explicit inclusion of district reduced RMSE from 0.534 to 0.512 and increased R^2 from 0.8659 to 0.8766, corresponding to $\Delta R^2 = +0.0107$. The gain was small but shows that there was other information in district membership that was not captured by the measured characteristics of the beneficiaries and programmes.

The benefit-adequacy ablation, on the other hand, showed negligible incremental predictive value. With the inclusion of perceived benefit adequacy, RMSE decreased from 0.535 to 0.534 and R^2 increased from 0.8653 to 0.8659, which is only $\Delta R^2 = +0.0006$. Therefore, perceived benefit adequacy appeared to have moderate SHAP attribution for a few predictions but contributed little to the overall predictive ability when demographic and accessibility variables were already accounted for.

Overall, the results suggest that PMMV accessibility is the most important predictive dimension of maternal health behaviour in the pooled sample, and that the ease of enrolment is the most important single dimension associated with digitalisation. The proposed Gradient Boosting model performed better than the traditional benchmark model and was relatively robust when predictors were grouped by predictor-profile. However, there is significant contextual dependence with high district heterogeneity and low cross-district transfer. The pooled R^2 of 0.865 is therefore not to be interpreted as proof of the external validity of the predictive framework: it is proof of the internal validity of the predictive framework within the combined analytical sample.

5. Discussion

The results show that the nonlinear Gradient Boosting model outperformed the top conventional model in predicting maternal health behavior for PMMVY beneficiaries. The proposed model resulted in R2 of 0.865 and RMSE of 0.535 which showed that the prediction error of the model is less than Ridge regression. This enhancement indicates that nonlinear predictive methods are better suited to capture the nonlinear associations between the demographic factors, accessibility measures, and maternal health behaviors. Gradient Boosting is especially well adapted to modelling nonlinearities, threshold effects, and interactions without having to specify them in advance. Advanced predictive methods have also been demonstrated to be useful for predicting the uptake of ANC and to identify key determinants beyond traditional statistical models (Sani et al., 2025; Osborne et al., 2026).

The most important predictive dimension was accessibility. Digitalisation related ease of enrolment was the most important feature in both Permutation Feature Importance and SHAP, followed by ease of documentation and ASHA/Anganwadi assistance. This aligns with the evidence that geographic and service-access factors influence maternal and newborn healthcare utilisation in rural India (Tripathi et al., 2024). They also align with evidence of the effectiveness of digitally-supported health-system interventions to enhance maternal and child health system utilization (Sharda et al., 2026). Administrative simplicity, digital accessibility and good support for frontline workers thus seem to be key aspects of programme accessibility.

There were also some differences between the PFI and SHAP rankings. The difference between PFI and SHAP is that PFI measures the decrease in predictive performance when a feature is disrupted, while SHAP measures the contribution of a feature to a fitted prediction for an individual. Thus, a predictor can affect specific predictions, but not significantly increase the accuracy of the model. This was the case for perceived benefit adequacy, which was the fifth in SHAP and seventh in PFI.

This finding was further confirmed by the benefit-adequacy ablation. Once the demographic and accessibility variables were added, the change in R^2 due to the addition of perceived adequacy was very small (0.0006). So, the results do not indicate that benefit adequacy is a good independent predictor in this data set. Rather, accessibility-related characteristics contained substantially greater predictive information. This does not necessarily be interpreted as a lack of policy importance of financial assistance, but rather that perceived adequacy did not provide much predictive value within the current specification.

A significant variation in districts was also found. The cross-district validation yielded very negative R^2 values, indicating that the relationships learned in one district were not very effective in the other. This finding aligns with the geographic variation in accessing maternal healthcare services in rural areas reported in the literature (Tripathi et al., 2024). Therefore, high pooled internal performance should not be considered indicative of geographic transportability.

The results from a programme perspective underline the importance of the usability of digital enrolment, the simplification of documentation and the assistance of frontline workers. But, all relationships identified are predictive and not causal. Key limitations are that the study is cross-sectional, is limited to two districts, has a maximum concentration in MHB, and has high district-specific distributions of accessibility and benefit adequacy. Future research should include larger multi-district samples, longitudinal designs, external validation, and ordinal and ceiling-sensitive modelling approaches to test for stability of these predictive patterns across more varied PMMVY settings.

6. Conclusion

The present study aimed to build an explainable predictive model to gauge the maternal health behaviour of the PMMVY beneficiaries in Karauli and Dholpur districts of Rajasthan. The study explored the predictive contribution of perceived adequacy of PMMVY benefits, programme accessibility, and demographic characteristics using an analytical sample of 350 rural beneficiaries. The proposed Gradient Boosting model showed good pooled internal predictive performance with an R2 of 0.865 and an RMSE of 0.535, and was superior to the best conventional Ridge regression benchmark. Perceived benefit adequacy was a much less important source of predictive information than accessibility-related factors. Under both Permutation Feature Importance and SHAP analysis, ease of enrolment related to digitalisation proved to be the most influential predictor. Documentation ease and frontline-worker assistance also made a significant contribution to prediction. By contrast, perceived benefit adequacy yielded only a marginal incremental gain in model fit, suggesting that the benefit of this variable to the model in its current form was relatively small. There was significant district variation and low cross-district transportability in the robustness analyses. The pooled model was well validated internally, but relationships found in one district were not well generalised to the other district. These findings highlight the need to consider local implementation context in interpreting predictive performance. The study presents a nested model selection, feature attribution, ablation analysis, and geographic transportability testing explainable predictive framework overall. But the framework is not ready for large-scale deployment without external validation. Future studies should use larger multi-district samples, longitudinal data, and ceiling-sensitive or ordinal modelling approaches to enhance generalisability and practical application.

Acknowledgment

This research paper is part of the project titled "Assessing Maternal Health Behavior under Heterogeneity Pradhan Mantri Matru Vandana Yojana (PMMVY) in North-Eastern Rajasthan: An Explainable Predictive Analysis of Accessibility, Benefit Adequacy, and District Heterogeneity", funded by the Indian Council of Social Science Research (ICSSR) under the Special Call for "Punyashloka Devi Ahilyabai Holkar Scheme for Women-Led Research on Women-Led Development."

References

1. Mehta, B. S., Alambusha, R., Misra, A., Mehta, N., & Madan, A. (2023). Assessment of utilisation of government programmes and services by pregnant women in India. PLOS ONE, 18(10), e0285715. doi:10.1371/journal.pone.0285715
2. Chi, H., Jung, S., Subramanian, S. V., & Kim, R. (2024). Socioeconomic and geographic inequalities in antenatal and postnatal care components in India, 2016–2021. Scientific Reports, 14, 10221. doi:10.1038/s41598-024-59981-w
3. Longchar, W., Kodali, P. B., & Hense, S. (2025). Trends and determinants of maternal health services utilization in India from 2015 to 2021. Scientific Reports, 15, 3711. doi:10.1038/s41598-025-87975-9
4. Banerjee, R., Singh, B. S., Sahrawat, N., Mehrotra, A., Swain, S., Mishra, R. M., Antony, P., Aparna, U., Jain, B. K., & Neogi, S. B. (2025). Effect of India's flagship conditional maternity benefit scheme on utilization of maternal and child health services: Evidence from a statewide evaluation study. BMC Pregnancy and Childbirth, 25, 294. doi:10.1186/s12884-025-07416-3
5. Jagannath, R., & Chakravarthy, V. (2025). The impact of Pradhan Mantri Matru Vandana Yojna scheme on access to services among mothers and children and their improved health and nutritional outcomes. Frontiers in Nutrition, 11, 1513815. doi:10.3389/fnut.2024.1513815
6. Ray, S., Nguyen, P. H., Pal, S., Scott, S., Menon, P., & Chakrabarti, S. (2026). India's national maternity benefit cash transfer program and child anthropometry. Scientific Reports, 16, 18807. doi:10.1038/s41598-026-48160-8
7. Marchant, T., Bottomley, C., Sahota, R., Porwal, A., Wadhwa, N., Chanda, S., Ranjan, R., Santhanam, D., Soni, M., & Das, A. (2026). Effects of a cash-plus intervention combining conditional cash transfer with social and behaviour-change communication on pregnancy weight and birth weight in India. Nature Health. doi:10.1038/s44360-026-00143-9
8. Mishra, N., & Vellakkal, S. (2026). Conditional cash incentives, community health workers, and continuum of maternal and child healthcare: Evidence from India. Health Policy and Planning, 41(4), 612–626. doi:10.1093/heapol/czag019
9. Mathur, A., & Sen, G. (2025). Money matters: Evidence from a conditional cash transfer scheme on child health. Journal of Human Capital, 19(3), 435–473. doi:10.1086/734859
10. Casebolt, M. T., Singh, K., Speizer, I. S., & Halpern, C. T. (2023). Maternal healthcare use by women with disabilities in Rajasthan, India: A secondary analysis of the Annual Health Survey. Maternal Health, Neonatology and Perinatology, 9, 11. doi:10.1186/s40748-023-00165-1
11. Girotra, S., Malik, M., Roy, S., & Basu, S. (2023). Utilization and determinants of adequate quality

antenatal care services in India: Evidence from the National Family Health Survey (NFHS-5) (2019–21). BMC Pregnancy and Childbirth, 23, 800. doi:10.1186/s12884-023-06117-z

12. Tripathy, A., & Mishra, P. S. (2023). Inequality in time to first antenatal care visits and its predictors among pregnant women in India: An evidence from National Family Health Survey. Scientific Reports, 13, 4706. doi:10.1038/s41598-023-31902-3
13. Kekre, A., & Mahajan, K. (2023). Maternity support and child health: Unintended gendered effects. Journal of Comparative Economics, 51(3), 880–898. doi:10.1016/j.jce.2023.03.002
14. Manna, S., Singh, D., Ghosal, S., Rehman, T., Kanungo, S., & Pati, S. (2023). Out-of-pocket expenditure and its correlates for institutional deliveries in private and public healthcare sectors in India: Findings from NFHS-5. BMC Public Health, 23, 1474. doi:10.1186/s12889-023-16352-w
15. Singh, R. R., Sharma, A., & Mohanty, S. K. (2023). Out of pocket expenditure and distress financing on cesarean delivery in India: Evidence from NFHS-5. BMC Health Services Research, 23, 966. doi:10.1186/s12913-023-09980-w
16. Dandona, R., Majumder, M., & Kumar, G. A. (2024). Population-level trends over a decade in geographical inequality for opportunity in access to maternal care services: A cross-sectional analysis from the National Family Health Surveys in India. BMJ Open, 14(11), e083922. doi:10.1136/bmjopen-2024-083922
17. Tripathi, P., Chakrabarty, M., Singh, A., & Let, S. (2024). Geographic disparities and determinants of full utilization of the continuum of maternal and newborn healthcare services in rural India. BMC Public Health, 24, 3378. doi:10.1186/s12889-024-20714-3
18. Ghosal, J., Nanda, A., Behera, P., Som, M., Bal, M., Ranjit, M., Kanungo, S., Kadam, S. M., Pati, S., & Dutta, A. (2025). Coverage of continuum of Maternal, Newborn and Child Health (MNCH) care in a nationally representative sample of 40,687 mother-child dyads of India: A report from NFHS-5. BMC Public Health, 25, 2727. doi:10.1186/s12889-025-23697-x
19. Singh, P., Singh, S., & Singh, K. K. (2025). Pregnancy to postpartum: Analyzing dropouts and socioeconomic predictors of continuum of maternal healthcare in India. Dialogues in Health, 7, 100226. doi:10.1016/j.dialog.2025.100226
20. Sharda, S., Singh, P., Trakroo, A., Agrawal, P. K., Goyal, A., Agarwal, A., Chawla, M., Joshi, A., Singh, S., Singh, R., Kshtriya, P., D'Aquino, L., Kumar, S., & Prinja, S. (2026). Impact of digital supportive supervision (DiSS) on the extent of maternal and child healthcare service utilisation in India: A sequential mixed-methods quasi-experimental study. BMJ Open, 16(1), e099539. doi:10.1136/bmjopen-2025-099539
21. Sani, J., Oluwagbemiga, A., & Ahmed, M. M. (2025). Machine learning-based prediction of optimal antenatal care utilization among reproductive women in Nigeria. Machine Learning with Applications, 21, 100698. doi:10.1016/j.mlwa.2025.100698
22. Osborne, A., Olawade, D. B., Oton, E. A., Odey, L. M., & Usani, K. (2026). Hierarchical machine learning models for predicting antenatal care utilisation among Nigerian women: Identifying actionable insights for health policy. BioData Mining, 19(1), 24. doi:10.1186/s13040-026-00538-0
23. Kassaw, A. K., Kidie, A. A., Tesfa, N. A., Mislu, E., Yimer, A., Hailu, M., Yesuf, H. A., Ayal, B. G., Geto, A. K., & Feleke, S. F. (2025). Optimizing machine learning models for predicting health service access and determinants among pregnant women in rural Ethiopia. Scientific Reports, 15, 40559. doi:10.1038/s41598-025-24245-8
24. Noor, S. T. A., Islam, R. B., Yeasar, S., & Siddique, S. (2026). Machine learning-based prediction of maternal continuum of care completion: Evidence from Bangladesh Demographic and Health Survey 2022. Array, 29, 100666. doi:10.1016/j.array.2025.100666
25. Dhar, A. K., Akther, S., & Bina, F. A. (2026). Determinants of maternal postnatal care utilization in Bangladesh: A machine learning and SHAP-based analysis of BDHS 2022 data. PLOS ONE, 21(5), e0350188. doi:10.1371/journal.pone.0350188
26. Kebede, S. D., Walle, A. D., Mamo, D. N., Enyew, E. B., Adem, J. B., & Alemayehu, M. A. (2025). Leveraging machine learning to identify determinants of zero utilization of maternal continuum of care in Ethiopia: Insights from SHAP analysis and the 2019 Mini DHS. PLOS Global Public Health, 5(6), e0004787. doi:10.1371/journal.pgph.0004787
27. Yoseph, A., Mussie, L., & Belayneh, M. (2026). Predicting antenatal care dropout using machine learning and describing maternal and neonatal outcomes by dropout status in Sidama Region, Ethiopia: A retrospective cohort study. BMJ Open, 16(7), e111423. doi:10.1136/bmjopen-2025-111423.