Vol. 06, No. 02, pp. 169 –182 (2025)
ISSN: 2708-0757

JOURNAL OF APPLIED SCIENCE AND TECHNOLOGY TRENDS

www.jastt.org

Feature-Based Child Mortality Prediction Using Ensemble and Traditional Machine Learning Models

Sampathirao Aruna*, Modalavalasa Divya, Promod Kumar Sahu

Department of Computer Science and Engineering, Aditya Institute of Technology and Management, Tekkali, Srikakulam Andhra Pradesh, India. arunamtech09@gmail.com,divya.modalavalasa@gmail.com, pksahu@adityatekkali.edu.in

*Corresponding author: arunamtech09@gmail.com

Abstract

Child mortality is a big problem around the world, especially in low- and middle-income nations where there are big differences in health care and social conditions. This investigation seeks to create a predictive model for child mortality and pinpoint the key factors that significantly contribute to it, employing machine learning (ML) methodologies. The dataset includes various features such as parental age, maternal education, birth weight, wealth index, and access to healthcare services. Thirteen machine learning classifiers were used, categorized into four model groups: Traditional Models (Logistic Regression, K-Nearest Neighbors, Support Vector Machine, Naive Bayes), Tree-Based Models (Decision Tree, Random Forest, Extra Trees), Boosting Models (AdaBoost, Gradient Boosting, XGBoost), and Ensemble Learning Models (Soft Voting, Hard Voting, Stacking). The efficacy of each model was assessed using classification metrics including Accuracy, Precision, Recall, and F1-Score within a 10-fold cross-validation framework to guarantee robustness. Results indicate that ensemble models, particularly AdaBoost, achieved the highest predictive accuracy, with perfect scores across all metrics (1.00). XGBoost and Stacking also demonstrated strong and consistent performance. The findings indicate that ensemble learning methods are effective in predicting child mortality and can assist policymakers and healthcare planners in identifying high-risk populations and implementing targeted interventions to reduce child mortality.

Keywords: *Child Mortality Prediction, Machine Learning Classifiers, Ensemble Learning Models, Explainable AI (XAI), Healthcare Data Analytics*

Received: June 14th, 2025 / Revised: July 30th, 2025 / Accepted: August 6th, 2025 / Online: August 8th, 2025

I. INTRODUCTION

Child mortality is a serious public health issue, especially in low-income and developing nations where there are significant differences in mother education, socioeconomic status, and healthcare access [1,2]. The World Health Organization (WHO) reports that 4.8 million children under the age of five die annually from preventable or treatable causes, including inadequate prenatal care, insufficient immunization, malnutrition, and limited access to clean water and sanitation [3]. Reducing child mortality is not only a moral imperative but also a key indicator of a nation's overall development and healthcare effectiveness. Understanding the multifaceted causes and associated risk patterns of child mortality is vital for designing effective interventions and evidence-based policy-making [4].

In this context, technologies on the data such as Machine Learning (ML) have become known as powerful tools in healthcare analytics. Unlike traditional statistical methods, ML algorithms are capable of acquiring hidden patterns and representing complicated nonlinear relationships among multiple features gives powerful potentials for early identification of children at risk of mortality [5]. Machine learning creates new techniques to use prediction models in health care like predicting child mortality in the recent days. By taking large and diverse datasets, ML models can be used to develop robust, data-driven decision support systems. These systems can assistance policymakers and healthcare professionals in identifying liable people and providing needs more effectively.

In this study, we have used a complex dataset that includes a wide range of demographic, health, and socio-economic features such as maternal and paternal age, birth order, birth weight, maternal education, wealth index, prenatal check-ups, institutional delivery, vaccination status, and access to water and sanitation. These features were carefully selected based on their known associations with child mortality outcomes. To evaluate the predictive capability of machine learning in this domain, the

doi: 10.38094/jastt62264

study systematically analyzes and compares 13 different classifiers, including traditional models (Support Vector Machine, Logistic Regression, K-Nearest Neighbors, Naive Bayes), tree-based models (Decision Tree, Random Forest), boosting algorithms (AdaBoost, Gradient Boosting, XGBoost), and ensemble methods (Soft Voting, Hard Voting, and Stacking). Each classifier is trained and tested on the given dataset, and their performance is calculated using evaluation metrics. The study integrates predictive modeling with Explainable AI (XAI) techniques, particularly utilizing SHAP (SHapley Additive Explanations) for the interpretation of model predictions. This allows us to identify which features contribute most significantly to child mortality risk, thus adding a layer of transparency and interpretability to the machine learning outputs. This research makes three major contributions: first, it presents a comprehensive benchmarking of 16 machine learning models on a real-world child mortality dataset; second, it offers a detailed comparative analysis of traditional and ensemble models for child mortality prediction; and third, it applies SHAP for feature importance analysis to enhance the interpretability and reliability of the model outcomes.

II. Literature Survey

F. Iqbal et al.,[6] discussed how child mortality prediction is possible using predictive analytics. The author has considered the SMOTE technique to balance the dataset. Apart from this, they also used 4 supervised machine learning classifiers for the said task. They employed performance metrics for comparative analysis to predict the child's survival status. The number of children under five in a household, the time between pregnancies, the size of the family, the mother's age at her first birth, antenatal care visits, breastfeeding practices, the child's birth weight, and the place of delivery are all important risk factors for child mortality. The random forest classifier was able to predict the deaths of children under five with an accuracy of 93.8%. The findings may substantially enhance the decision-making process for child health intervention initiatives.

A. W. Demsash et al.,[7] investigated machine learning classifiers for child mortality prediction. The author has considered the five algorithms, like J48, RF, etc. They used 1813 samples for testing child mortality from the 2019 Ethiopian Demographic and Health Survey dataset. The author has divided the entire dataset into 2 parts,70% training and 30% testing. Each algorithm has undergone 10-fold cross- validation. For generating the rule, authors were considered if…then rules and implemented through WEKA Software.

C. Chivardi et al. [8] used machine learning to study how socioeconomic factors affected child mortality rates (U5MR) in Brazil, Ecuador, and Mexico over a period of 20 years. We built a cohort at the municipal level from 2000 to 2019 and trained a random forest model to see how important socioeconomic factors are in predicting U5MR. As part of a sensitivity investigation, we trained two more machine learning models and reported the mean squared error, median absolute deviation, and root mean squared error. According to our statistics, the Gini coefficient, poverty, and illiteracy are the best indicators of U5MR.

O. Samuel et al.,[9] studied machine learning models and discussed how child mortality is possible as well as how to predict under-five mortality, with the Random Forest and ANN algorithms. This model obtained an accuracy of 89.47% and an AUROC of 96%. It has been observed that under-five mortality rates have some of the features, influenced more by wealth index, maternal education, antenatal visits, etc, in Nigeria.

To determine the effect of overall, health-related, and other types of public spending on child mortality, L. P. Garcia et al.[10] examined data from 147 nations (2012–2019). Increased public spending, especially in health, was associated with lower neonatal mortality, while investment in non-health sectors reduced death in children aged 28 days to 5 years, according to results obtained using a Generalized Propensity Score approach. It is clear that the amount and distribution of public investments are crucial in enhancing child survival, since the effects differed by industry and age group.

S. Holcroft et al. [11] aim to construct a PPH prediction model by examining the statistical importance of early risk factors. The data set originated from an observational research in northern Rwanda that used a case-control design. Random Forests, Extremely Randomized Trees, logistic regression, logistic regression with elastic-net regularization, gradient-boosted trees with XGBoost, and logistic regression were the machine learning methodologies and statistical models that were examined. Since the Random Forest model had better results than the other models (average sensitivity: 80.7%, specificity: 71.3%, and misclassification rate: 12.19 percent), it may be a useful tool for PPH prediction.

R. K. Saroj et al.,[12] took the data from the Family Health Survey (NFHS-IV) to predict child mortality using a machine learning classifier. The author has considered DT, RF, NB, KNN, LR, etc. Each model's performance metric was estimated, followed by a confusion matrix. Among the predictive models included in this study, the neural network model proved to be the most effective in predicting death among children under five. The most effective model was the neural network, while logistic regression also performed well in terms of accuracy (94% to 95%), precision- recall curve (99.5% to 99.6%), and AUROC range (93.4% to 94.8%) in predicting under-five mortality.

A. Satty et al.[13] look at how machine learning classifiers use data from the 2018–2019 Multiple Indicator Cluster Survey to estimate the death rate of children under five in CAR. The author used LGB, XGBoost, EGB, and compared them with the traditional logistic regression model. As well as evaluated the performance metrics and found that CATBoost is one of the top performers for child mortality prediction with an accuracy of 0.94, F1-score is 0.95, and AUC is 0.973.

S. Naznin et al. [14] looked into how well a machine learning model could predict child deaths. The author has looked at the information from the Bangladesh Demographic and Health Survey. The under-five mortality rate in Bangladesh has markedly declined during the research period, and machine learning models effectively predict future trends. Linear Regression demonstrated the highest accuracy among the models, evidenced by the lowest Mean Absolute Error (4.05), Root Mean Square Error (4.56), and Mean Absolute Percentage Error (6.64%), as well as the highest R-squared value (0.98).

S. Das et al.,[15] investigated the collected data from Bangladesh with 5669 children. The study looked at the medical records of kids aged 0 to 59 months who were hospitalized to the acute care unit at the International Centre for Diarrheal Disease Research in Dhaka, Bangladesh. We tested logistic regression, gradient boosting trees, random forest, elastic net, least absolute shrinkage and selection operator, and our other models by looking at the area under the receiver operating characteristic curve.

S. Adithya et al., [16] focuses on predicting child mortality, specifically for children under the age of five, including fetal deaths. It aims to develop AI-based strategies to accurately determine factors affecting fetal and child well-being. The approach involves analyzing the dataset using supervised machine learning techniques, which includes steps like identifying variables, handling missing values, and performing univariate, bivariate, and multivariate analysis, along with data cleaning, validation, and visualization. The study also includes sensitivity analysis to understand how different model parameters impact fetal health classification. Finally, the study gives a machine learning-based framework for predicting child mortality and tests how well different ML models operate on the dataset.

C. Ashwini et al. [17] utilized the NFHS-V dataset to analyze spatial disparities in under-five mortality in Uttar Pradesh. Four machine learning models were employed to ascertain significant causes of mortality. Prediction accuracies varied between 76% and 79.4%, with logistic regression attaining the best accuracy, highlighting notable geographical variations. The factors that notably influenced the mortality rate of children under five years old encompassed the mother's body mass index (BMI), the number of births in the preceding five years, the child's gender, the timing of birth, prenatal therapy, birth order, and water accessibility. Machine learning approaches, especially logistic regression, can effectively influence actions aimed at enhancing child survival based on the findings.

By using machine learning to the 1,188 cases included in the publicly available "Paediatric Intensive Care database," J. Prithula et al. [18] hopes to enhance the accuracy of pediatric intensive care unit death prediction. Using Random Forest, Extra Trees, and XGBoost, we selected 16 important characteristics from 105. Then, we tested 10 ML models, including CatBoost and ensemble approaches. Addressing class imbalance, a novel data splitting strategy dramatically improved performance. The use of ML to improve intensive care unit readiness and clinical outcomes was demonstrated by the 85.2% AUC and 89.32% accuracy achieved by the suggested strategy.

Along the Iraq–Turkey border in the Duhok region, the study [19] examines forest change from 2015–2024. It used machine learning and satellite imagery to show that forest loss has been on the rise, mostly as a result of human activity like road construction, fires, and illicit deforestation. There was a reduction in forest cover from 630 km^2 to 577 km^2. Results from the tests showed that XGBoost was the most effective model. This research demonstrates how machine learning may fill in data gaps in our understanding of forest change and its management.

A. A. Abdullah et al. [20] presents a novel approach to medical picture classification uncertainty assessment by utilizing Bayesian deep learning ensembles. This method improves prediction accuracy and confidence by selecting the top 'k' models according to their strength in predicting each class. This approach is helpful for high-risk domains like healthcare because it outperforms or performs comparably to traditional Bayesian ensembles.

In order to fix the overconfidence in predictions made by conventional deep learning models in healthcare, the study [21] examines how Bayesian Deep Learning (BDL) might be used. For delicate medical judgments, BDL is preferable than classical models since it assesses uncertainty. Medical imaging, clinical signals, and electronic health records are just a few of the fields that benefit from BDL methods, which are discussed in this article along with their limits. It goes on to mention current research obstacles and voids in healthcare BDL applications and goes over new DL designs.

The study by M.S Rao et al., [22] delves into the ways in which AI can enhance healthcare by facilitating more precise and rapid prediction. The analysis of massive volumes of health data is made easier with the use of machine learning and cloud computing. The use of AI in conjunction with wireless sensors that monitor one's behavior and level of physical activity allows for the early diagnosis of diseases. With the help of these innovations, healthcare systems are becoming more intelligent and adaptable. The article goes on to talk about how AI may improve decision-making using environmental data and how it can assist reduce mistakes in medicinal therapies.

The goal of the team-based competition conducted by B. A. Sullivan et al. [23] was to estimate the probability of mortality in NICU patients using several machine learning algorithms. For this study, five groups of neonatologists used a dataset including more than six thousand NICU cases using models such as logistic regression, neural networks, and XGBoost. Although the audience had a preference for the sophisticated CNN model, logistic regression produced the highest AUC and was the most accurate. The results show that simpler models can sometimes beat more complicated ones, demonstrating the importance of data comprehension and model interpretability over complexity.

J. Lee et al. [24] aimed to enhance mortality prediction in preterm neonates (<32 weeks gestation) during hospitalization. Traditional approaches (CRIB-II and logistic regression) were compared to random forest. The random forest model, trained to flag data 6 hours before death as "worry," beat traditional models on 275 newborns. The random forest model predicted mortality better by incorporating clinical and physiological data than CRIB-II and logistic regression, which had AUC scores of 0.78 and 0.84.

Predicting the mortality of infants was the subject of eleven research articles that C. Mangold et al. [25] reviewed. Out of 434 studies that were evaluated, including 1.26 million babies, only those with 500 or more participants and postnatal data were considered. It was common practice to employ logistic regression, neural networks, and random forests. Five minutes to seven days following delivery, some versions utilized three to sixty-six attributes. Very few studies were checked and adjusted by experts outside the field. With sensitivity levels ranging from 63-60% and specificities from 78% to 99%, AUC values were between 58.3% and 97%. Although it made use of a large number of features, linear discriminant analysis yielded the best results. Additional research is necessary to confirm and use ML for healthcare prediction of newborn mortality, according to the review.

Using WHO health data, K. Pal et al. [26] employ machine learning to forecast the potential mortality rate of newborns in various nations. Over the course of several years, the researchers merged general health statistics with data on reasons of mortality, such as prematurity, infections, or birth injuries. After comparing various prediction models, they discovered that the Random Forest model produced the most accurate results (R^2 = 0.990), indicating very high accuracy. This demonstrates how machine learning can assist healthcare providers and governments in making more informed decisions regarding the care of newborns.

III. METHODOLOGY

The methodology illustrates a systematic architecture designed to predict child mortality using different ML and ensemble models. As discussed in the system architecture, the procedure begins with the gathering of child mortality data, followed by pre-processing to remove unwanted data and normalize the dataset. Transformed data is obtained from preprocessing and afterward categorized into training and testing sets. Various models are implemented to train the processed dataset. The trained models are subsequently assessed using the testing data, which is based on evaluation metrics. The final predicted output is derived from the most effective model.

Initially a dataset is to be taken to process the data which is to be tabular format in the combination of rows and columns. The dataset used in our study is publicly available and can be accessed by link: https://data.mendeley.com/datasets/cfwnrgd9jm/1. The dataset includes both maternal and demographic attributes such as mother's age, education, number of prenatal visits, previous child deaths, type of residence, income category, and more. These features are known to be linked with child mortality risk. The dataset is moderately imbalanced, with approximately Positive class (mortality) as 49.5% and Negative class (survival) as 50.5%. The dataset was split into 80% training and 20% testing sets The mentioned Figure 1 is the complete architecture for child mortality. Our target variable is mortality, and it consists of a yes or no type. That means it is a binary class classifier. During the data preprocessing stage, we handled the missing values and encoded the categorical features using one-hot and label encoding, followed by scaling. Our objective was to prepare clean and consistent data suitable for the model. In this phase, we identified the top features to build the model. Out of 13 features, we only considered 9 features that affect child mortality. After training the model (using Extra Trees, Random Forest, or XGBoost), feature importance scores were calculated, and top features such as birth-weight, institutional-delivery, mother_education, and antenatal_visits were chosen to increase model interpretability and reduce overfitting.

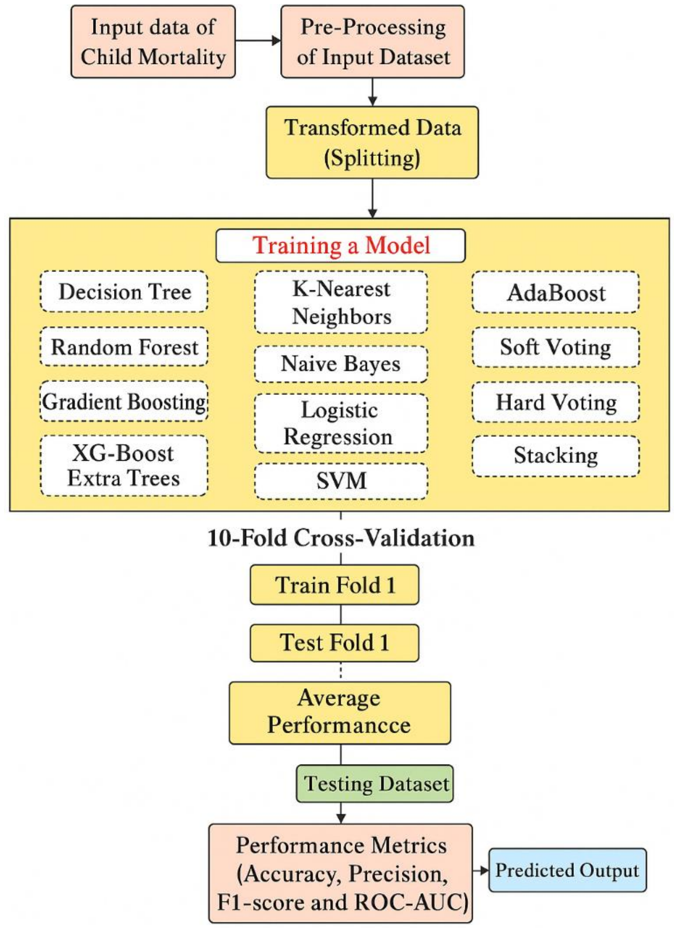


Fig. 1. Proposed Model for Child mortality classifier

A. Model selection

Different models have been used for child mortality, and these are Logistic Regression, Random Forest, KNN, Naive Bayes, Decision Tree, SVM, Gradient Boosting, XGBoost, Extra Trees, AdaBoost, Soft Voting, Hard Voting, and Stacking. Our goal is to compare the performance metrics and identify the best model. After training the model (e.g., Extra Trees, Random Forest, or XGBoost), feature importance scores were retrieved, and top features, including birth_weight, institutional_delivery, mother_education, and antenatal_visits, were chosen to increase model interpretability and reduce overfitting.

- *Logistic Regression:* It estimates the probability of a binary outcome in this case, whether a child is likely to survive or not. It performs well with linearly separable data and produces interpretable coefficients that makes understanding the effectiveness of each prediction as it is simple and efficient but may not capture non-linear relationships among features affecting child mortality.
- *Decision Tree:* Builds a hierarchical structure of decisions based on inputs like maternal education, birth weight, and access to healthcare. It splits the data into branches based on the correlated attributes, making it easy to interpret and visualize [27]. However, it leads to overfitting, probably with small variations in the dataset,

which may impact on generalization of unseen child health data. Information Gain helps in feature selection during tree construction [28].

Information Gain:

$$\mathrm{IG}(\mathrm{D},\mathrm{A}) = \mathrm{H}(\mathrm{D}) - \mathrm{H}(\mathrm{D}|\mathrm{A}) \tag{1}$$

where the entropy H(D) is given by

$$H(D) = -\sum_{i=1}^{n} P(x_i)\, log_2\big(P(x_i)\big) \tag{2}$$

Here $P(x_i)$ represents the probability of class x_i in dataset D, and H(D|A) is the conditional entropy after splitting on attribute A.

- *Random Forest:* Extended version of Decision Tree that elevates prediction accuracy and decreases overfitting by aggregating the outputs of multiple trees. It effectively identifies the feature correlations and handles missing values, making it suitable for complex health datasets [29].
- *K-Nearest Neighbors (KNN):* Predicts the results by taking the most similar data points (neighbors) in the feature space. In child mortality, it classifies based on points in closely related cases like similar maternal age and health access. However less effective with uncertain data.
- *Support Vector Machine (SVM):* Constructs a best possible hyperplane that separates the two classes (mortality vs. survival) in a high-dimensional space. It is particularly effective in handling complex boundaries and smaller datasets. In child mortality prediction, SVM helped model complicated mappings between socioeconomic and health indicators but it may require perfect kernel choice.
- *Naive Bayes:* It is probabilistic classifier depends on Bayes' theorem. Despite its simplicity, it performs well on categorical health data and can quickly highlight the most likely causes of mortality. It gives best results with high-dimensional datasets but its assumption of independence may not always be true in healthcare data.
- *Gradient Boosting:* Builds a sequential DT's where each new tree corrects the errors of the previous ones in order to boost the weak learning points at each previous trees [30]. It performs well on structured data and can model complex patterns but requires careful tuning to avoid overfitting.
- *XGBoost (Extreme Gradient Boosting):* It is a better variant of gradient boosting that is noted for being fast and efficient. It solves the problems of missing values, feature mappings and imbalancing. By using XGBoost in this research produced best results and highlighted key factors contributing to child deaths with high feature importance scores.
- *Extra Trees (Extremely Randomized Trees):* It constructs multiple trees using random splits and then aggregates the results. It adds more randomness so that it can reduce variance in future. It helped to identify correlations between features like institutional delivery and birth weight.
- *Soft Voting:* Similar to ensemble technique that averages the predicted probabilities of multiple base techniques to make a final decision. It benefits from the strength of each individual classifier and balancing the predictions in uncertain situations.
- *Hard Voting:* selects the majority class predicted by multiple classifiers. It is simple and effective when base models agree but may struggle if the models are diverse and inconsistent. It performed moderately well in this study by aggregating predictions from strong classifiers like random forest and SVM.
- *Stacking:* Advanced ensemble model were estimating of different base models are used as input to a meta-model, which makes the final speculation. It exploits the robustness of various algorithms and often improves predictive performance.
- *AdaBoost (Adaptive Boosting):* AdaBoost integrates multiple weak learners into a strong learners by concentrating more on difficult cases in each cycle. It regulates weights for misclassified data points and improves overall prediction accuracy. Among all models, AdaBoost showcased the best performance, making it the most powerful model for diagnose child mortality risk in the dataset.

B. Model Training and Validation

We split our data into two parts: training and testing. We utilized 80% of the training data to train the model and 20% of the data to test it. This step is used for learning the machine learning classifiers' patterns and generalizing the unseen data to predict the child mortality ((whether a child survived or not. Each model learns the mapping between the chosen input features and the target variable (mortality: 0 = alive, 1 = dead). Validation assures that the model does not overfit and remains generalizable for real-world use (for example, in public health decision-making). By evaluating 13 alternative algorithms, you may select the best-performing model that strikes a balance between high accuracy and strong sensitivity to mortality cases (minimizing false negatives).

C. Model Evaluation

During model evaluation, we used the performance metrics. This phase discusses how the model is trained and predicts child mortality. After the model is trained, each classifier is tested and gives information about whether a child will survive or not, based on the given features.

D. Prediction and Decision Making

In this phase, our model predicts the child mortality risk for unseen data. After being trained and verified using historical data with known outcomes, the model is used to estimate the chance of child mortality based on new input data (test or real-world). Our target Variable (Dependent Variable) is mortality, which contains value Binary value (0: Alive, 1: Deceased)

IV. Results And Discussions

Table I shows that evaluation metrics of 13 different ML and Ensemble Models and it has been observed that AdaBoost performed well in comparison to the other models. We used the 13 classifiers for child mortality prediction and estimated their performance [31]. Among them, the top performer is AdaBoost and which obtained a perfect score and it is 1.0. It means it works very well for classification on the test data. Apart from this classifier, stacking is also working well, and its accuracy is 0.997,f1 score is 0.989. Similarly, XGBoost is also performed, and its accuracy is 0.996, F1-score is 0.986. Both classifier shows very high recall and ROC-AUC. The model rate performers are RF, Extra Tree, and Hard Voting.

They exhibit high precision and good generalization. It also observed that the weaker performers for child mortality prediction are LR, NB, and KNN. But we observed that SVM performs a balanced model for child mortality prediction.

Evaluation Metrics: The evaluation metrics used in our study are [32-35]:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (5)$$

$$F1 - Score = 2 \times \frac{Precision \times \ Recall}{Precision \ + \ Recall} \quad (6)$$

TABLE I. PERFORMANCE METRICS OF DIFFERENT MACHINE LEARNING MODELS

S.No	Model	Precision	Recall	F1-Score	Accuracy	ROC AUC
1	Logistic Regression	0.7262	0.4326	0.5422	0.8970	0.9121
2	Decision Tree	0.9353	0.9220	0.9286	0.9800	0.9558
3	Random forest	0.9683	0.8652	0.9139	0.9770	0.9970
4	KNN	0.7064	0.5461	0.6160	0.9040	0.9072
5	SVM	0.8624	0.6667	0.7520	0.9380	0.9745
6	Naive Bayes	0.7750	0.4397	0.5611	0.9030	0.9372
7	Gradient Boosting	1.0000	0.9433	0.9708	0.9920	1.0000
8	XG-Boost	0.9858	0.9858	0.9858	0.9960	1.0000
9	Extra Trees	0.9565	0.7801	0.8594	0.9640	0.9929
10	AdaBoost	1.0000	1.0000	1.0000	1.0000	1.0000
11	Soft Voting	0.9712	0.9574	0.9643	0.9900	0.9983
12	Hard Voting	0.9601	0.9007	0.9304	0.9810	0.9474
13	Stacking	0.9790	0.9929	0.9859	0.9960	0.9999

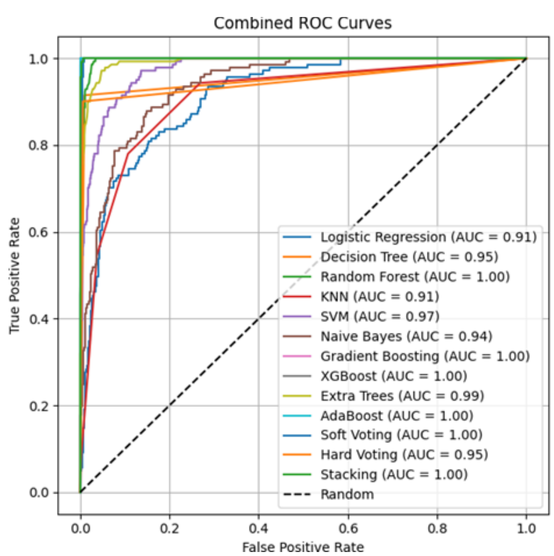


Fig. 2. Combined ROC Curve for child mortality prediction

The above-mentioned Figure 2 is the combined ROC curve. It provides the classification performance for child mortality prediction. He used 13 classifiers to distinguish the classes. In this figure, each curve represents the classifier, with the area under the curve representing the effectiveness of the model. It has been observed that models such as AdaBoost, Gradient Boosting, XGBoost, and Stacking perform well and achieve nearly perfect AUC values (≈1.0). We also observed that the other models like KNN and Logistic Regression showed lower AUCs, i.e, they are not well and accurate classification boundaries. The main diagonal line refers to the random performance. The line that is represented above the diagonal line is the best model and predictive capabilities.

Figure 3 is called the Precision-Recall curve for child mortality prediction. It provides the trade-off between precision and re-call for 13 classifiers. High-performing models such as AdaBoost, Stacking, and Gradient Boosting are perfect in situations where both false positives and false negatives are crucial, since they show big PR AUC scores in addition to maintaining high precision throughout a range of recall levels. Conversely, models like Naive Bayes and Logistic Regression show a smaller area under the curve, suggesting difficulties in successfully striking a balance between recall and precision. In unbalanced datasets, where ROC AUC might not accurately represent classifier performance, this visualization is especially helpful.

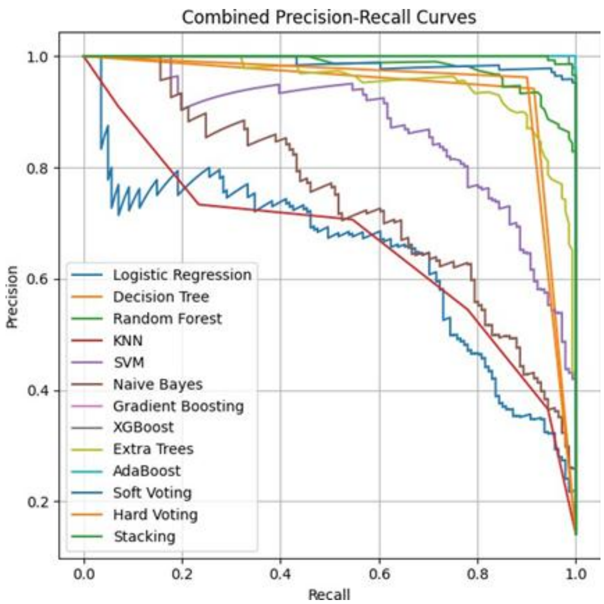


Fig. 3. Precision-recall curve for child mortality prediction

A. Logistic Regression

Figure 4 represents the logistic regression ROC curve and PR curve for child mortality prediction. We observed that the ROC curve is 0.9121, and accuracy is 0.8970, F1-score is 0.5422. It has been observed that there is good discrimination between classes. But the PR-AUC is relatively low We also observed that this model is not that much of suitable for identifying the correctly positive under class imbalance. The lower PR AUC suggests that it has trouble accurately detecting real child death instances under class imbalance.

B. Decision Tree

In Figure 5, the visualization is used to capture the non-linear relationships and interactions between features for child mortality prediction. It has been observed that the Decision Tree exhibited good performance, identifying significant patterns in the data with a ROC AUC of 0.9528 and PR AUC of 0.9111. Despite the possibility of overfitting, its comparatively high recall indicates that it successfully detects actual child death instances. The classifier captures the non-linear relationships and interactions between features. DT splits the data recursively and uses the required features to further build the model for estimation purposes.

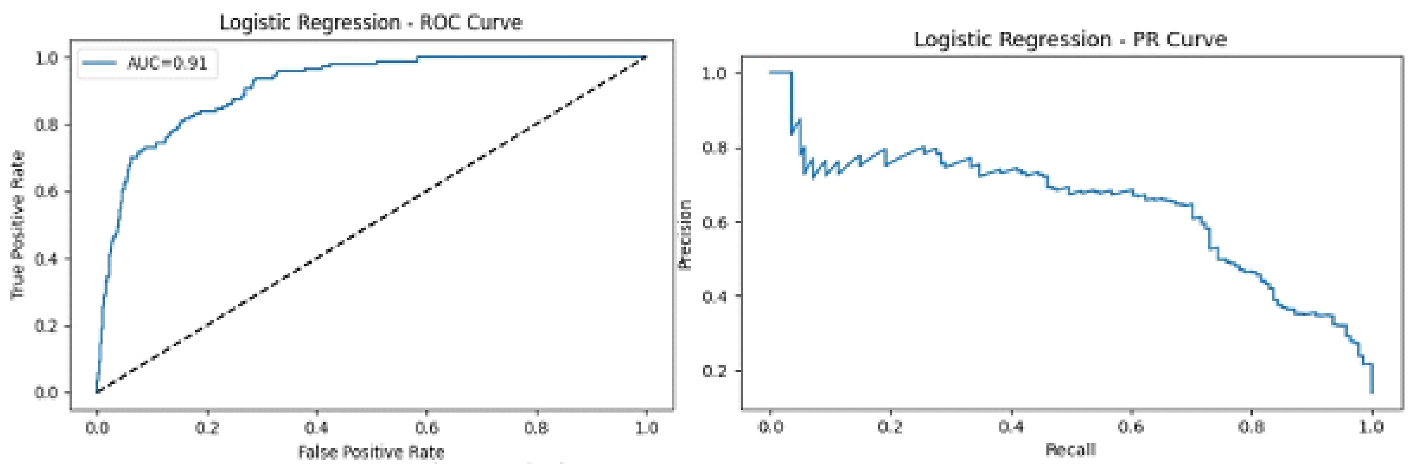


Fig. 4. ROC curve representation model for child mortality prediction.

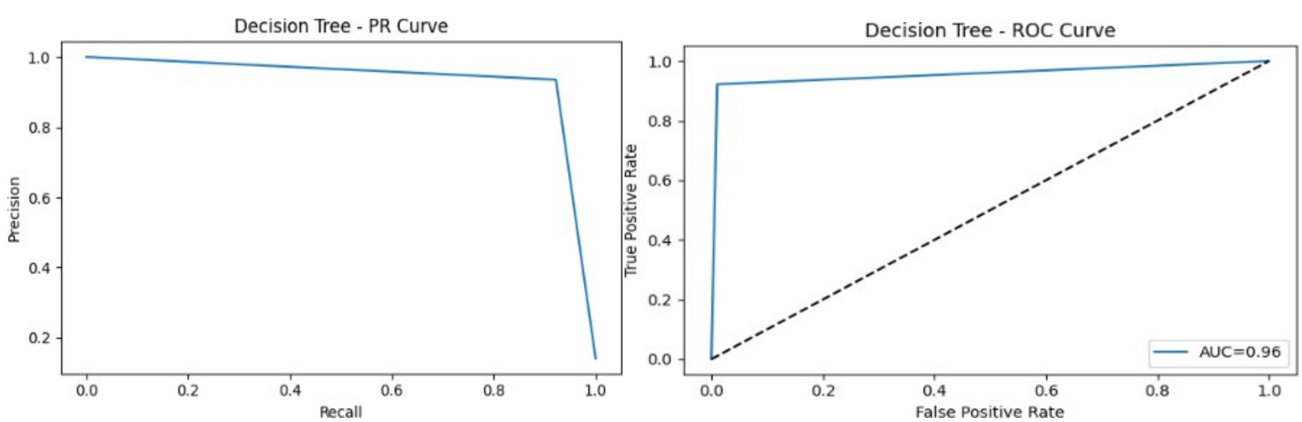


Fig. 5. ROC Curve and PR-Curve for Child Mortality Prediction

C. Random Forest

In the above-mentioned Figure 6, it is discussed how child mortality prediction is possible using the RF algorithm. It is used to reduce overfitting by averaging many decision trees. This model creates several trees with bootstrapped samples and averages their predictions. It has been observed that the ROC-AUC score is 0.997 and the PR-AUC is nearly 0.97 that the model is generalized and robust for child mortality prediction. It is among the best at classifying child mortality because of its ensemble nature, which enables it to represent intricate patterns and relationships. With a solid PR curve that maintains good precision across all recall levels and a ROC AUC of 0.9969, Random Forest stands out in the combined visualizations. This demonstrates how well the model can accurately detect cases of both positive and negative mortality with little overfitting.

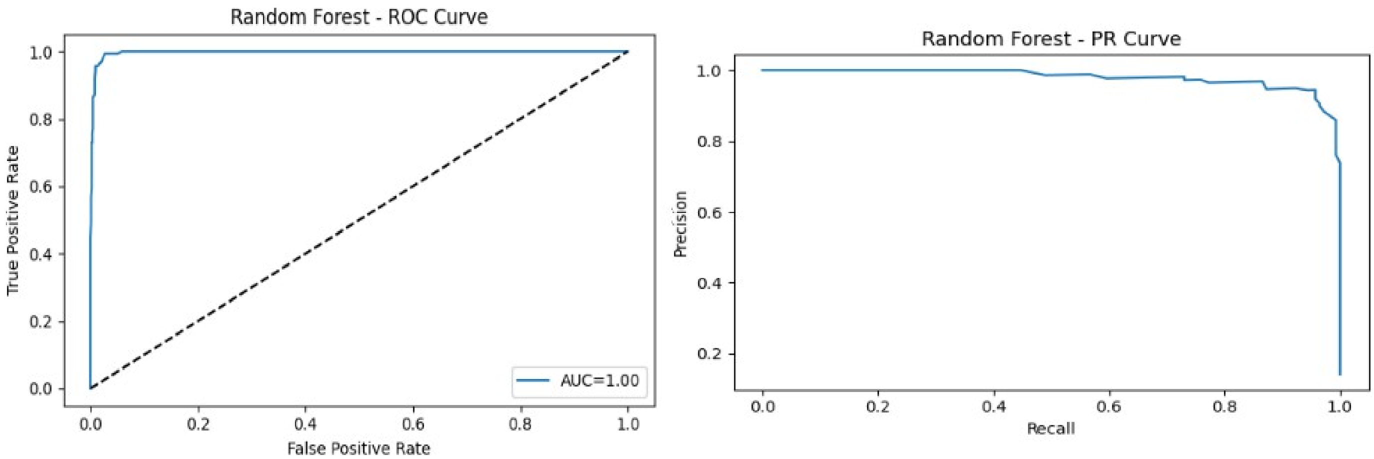


Fig. 6. Random Forest model for child mortality prediction

D. KNN

In the mentioned Figure 7 it visualized KNN-ROC and KNN PR-curve for child mortality prediction are visualized. The average discriminative capacity (AUC 0.907) is indicated by the moderately curved ROC curve for KNN. When attempting to record more real fatalities, the PR curve is noticeably flatter, particularly at higher recall levels, indicating that it frequently predicts false positives. This model is suitable for instance-based learning purposes, and it predicts the class of the closest training samples. This approach is less suitable for healthcare applications with delicate patterns since it is sensitive to irrelevant features and data scale.

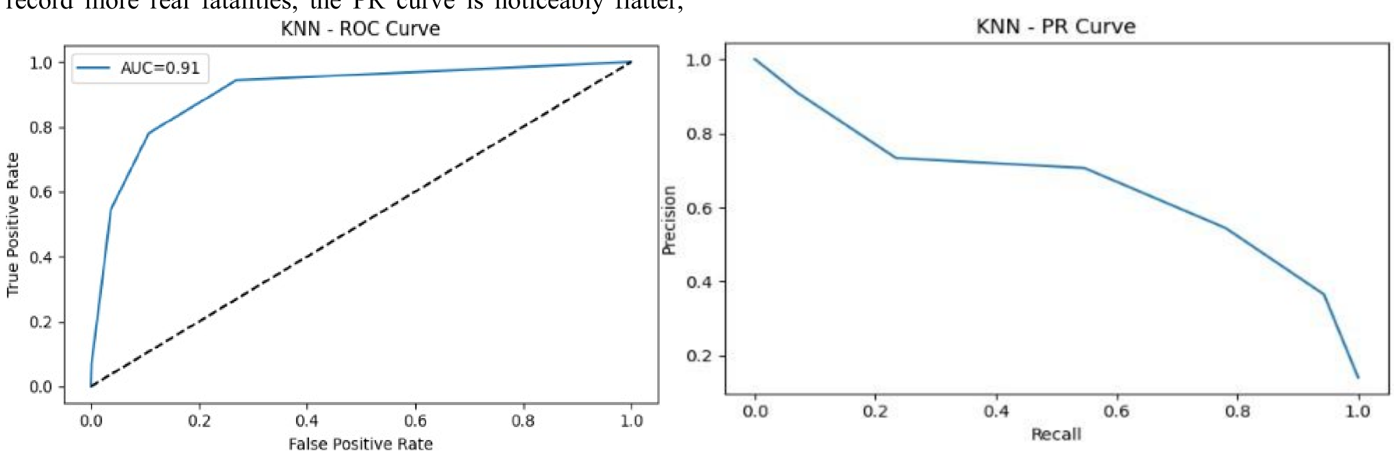


Fig. 7. KNN-ROC and PR curve for child mortality Prediction

E. SVM

Figure 8 discusses how SVM is useful for predicting child mortality. It has been observed that SVM has the strength to separate the classes using a hyperplane, and it obtained a ROC curve of 0.9745 and a PR-AUC is 0.8076. We used this model because it effectively handles the high-dimensional spaces with clear margins. It demonstrates the reliable performance, and it is suitable for mortality prediction. But it seems that this model suffers from being computationally expensive because it requires more tuning. Additionally, its PR curve is steady and continuously high, particularly at mid-to-high recall values, demonstrating its dependability in capturing more true positives while maintaining an acceptable level of precision.

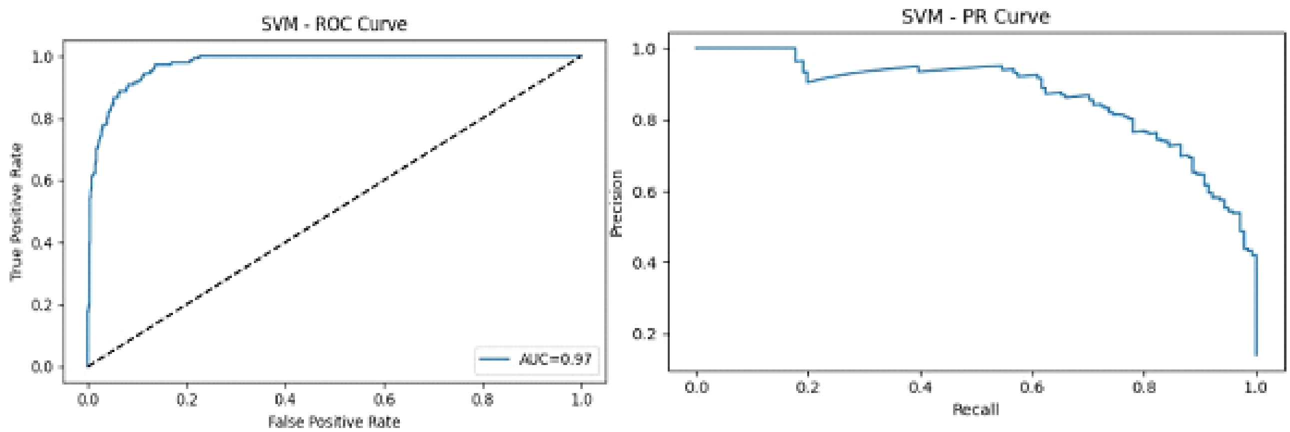


Fig. 8. SVM-ROC and PR-Curve for child mortality

F. Naive Bayes

The Figure 9 discusses how NB is suitable for child mortality prediction. It depicts the ROC Curve and PR curve. It has been observed that NB achieves a good ROC AUC score, and it is about 0.937, and the PR-curve score is higher for recall. In real-world situations with intricate data interdependencies, Naive Bayes is less reliable. Internally, this model estimates the conditional probabilities. The depicted model indicates that it often misclassifies positives.

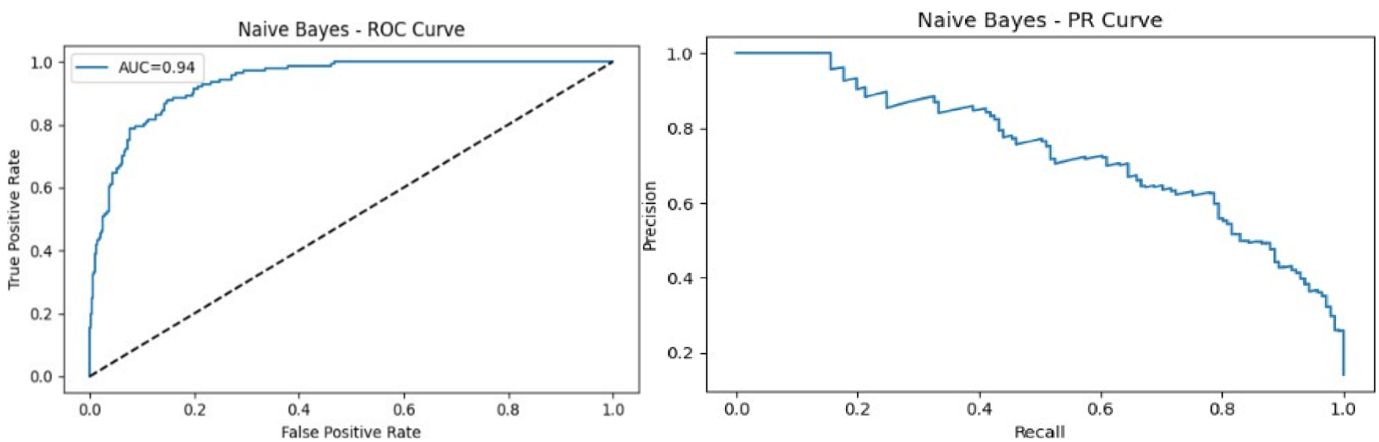


Fig. 9. NB_ROC and PR-curve for child mortality

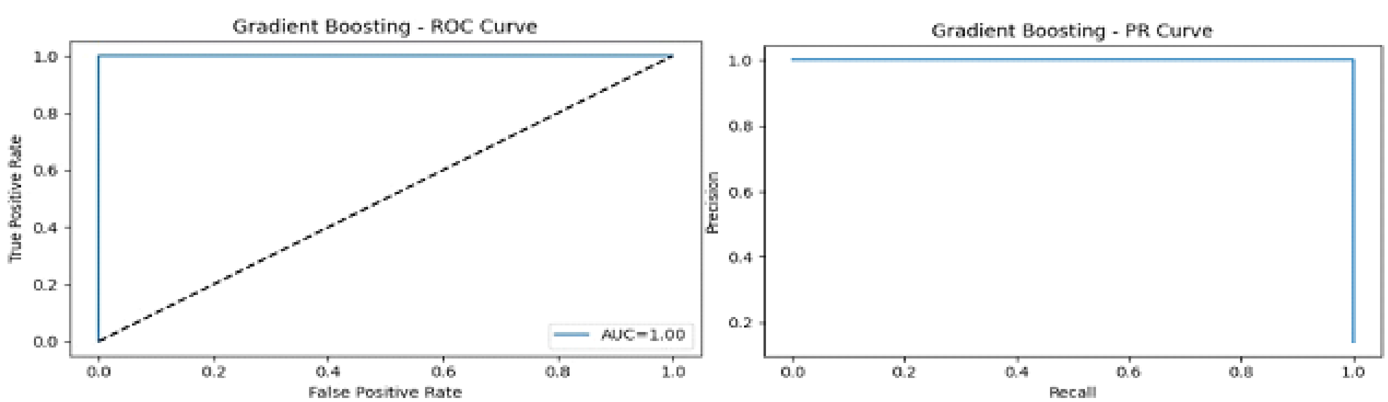


Fig. 10. GB-ROC and PR curve for child mortality

G. Gradient Boosting

Figure 10 represents the GB visualization for predicting child mortality. It discussed how the model learns iteratively to fix the errors and increase the overall performance. The PR-Curve score is nearly about to 0.97, and the ROC-AUC score is 1.0, which means this model makes precise and sensitive predictions. It is considered one of the good models for mortality prediction. It can learn complex patterns like multi-factor causes of death (e.g., malnutrition + disease + environmental factors) by repeatedly building trees and fixing past mistakes. The form of its PR curve demonstrates its capacity to reduce false positives while preserving high sensitivity.

H. XGBoost

The above-mentioned Figure11 Discusses how XGBoost performed for child mortality prediction. The obtained ROC-AUC score is 0.999, and the PR-curve is (0.98+), making it near-perfect. Here we used XGB to avoid the overfitting model through a combined GDB and regularization technique. Its accuracy and recall remain excellent throughout the visualisations, demonstrating its applicability for delicate prediction tasks such as forecasting child mortality. It is among the most dependable models in terms of statistical performance and visualisation. This model is performing good and is especially valuable when precision and recall are both critical.

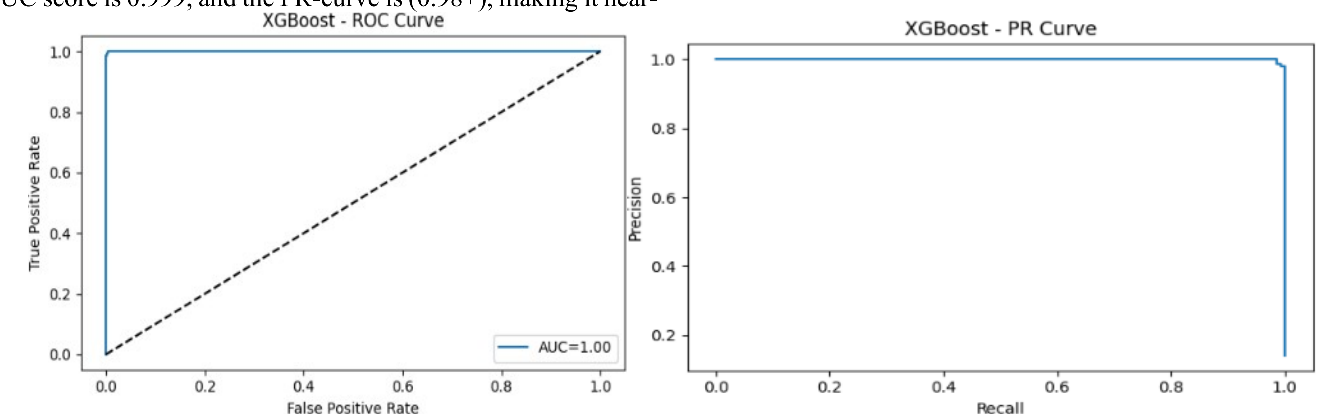


Fig. 11. XGBoost ROC-PR Curve for child mortality

I. Extra Trees

Figure 12 depicts extra tree classification for child mortality prediction. From the experimental observation, it is revealed that this model achieved a high ROC-AUC(0.995) and a strong PR curve (0.92).

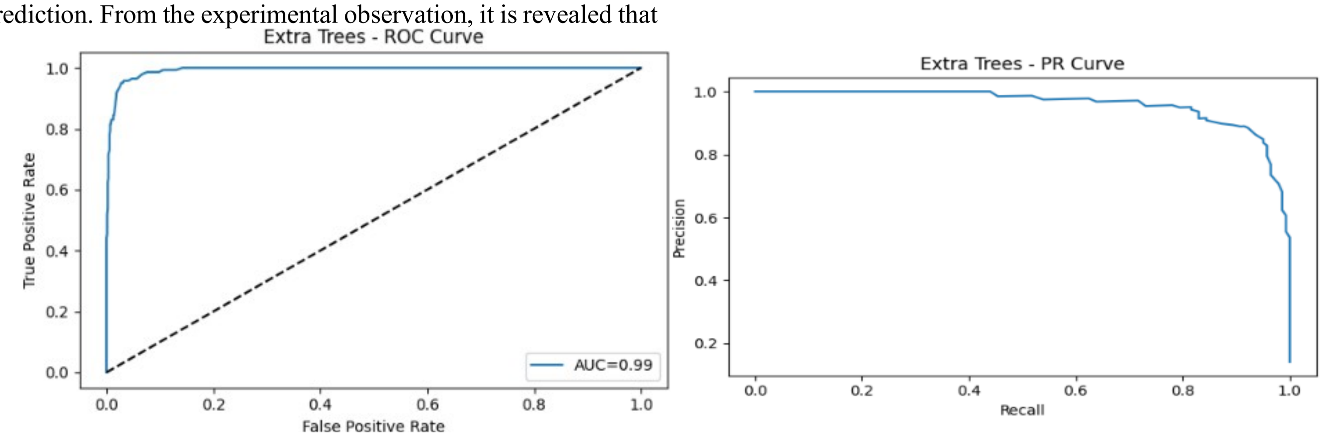


Fig. 12. ET_ROC and PR_Curve for child mortality

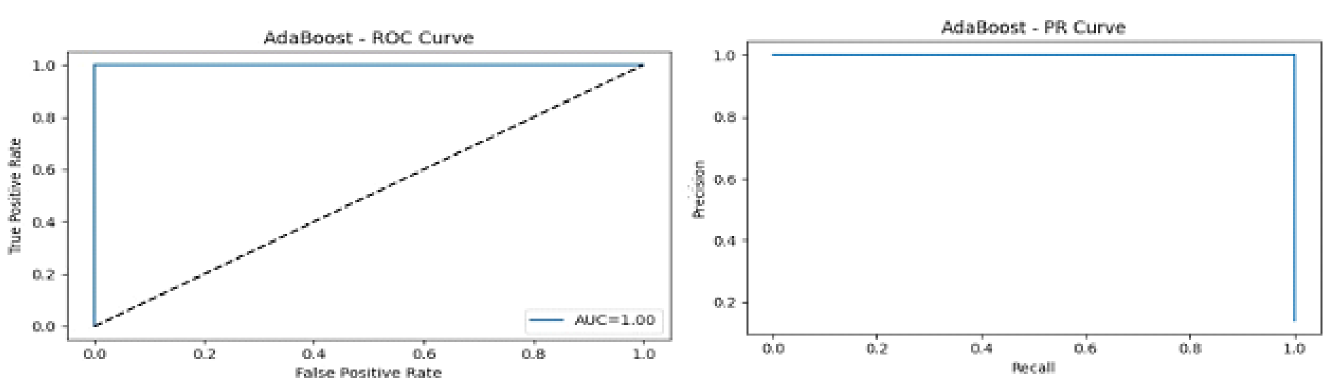


Fig. 13. AdaBoost_ROC and PR_Curve for child mortality

J. AdaBoost

Both ROC AUC (1.0000) and PR AUC (1.0000) were flawlessly achieved by AdaBoost, most likely as a result of its powerful boosting of weak classifiers. Although it might still be sensitive to noisy data, this makes it a unique model. This is the best model because it converts the weak learner into a strong classifier. Boosts misclassified samples to improve accuracy iteratively.

K. Soft Voting

It is one of the ensemble learning models that is used for child mortality prediction. It combines classifier probabilities for a more balanced output. Then it finds the average predicted probabilities for the different models, like RF, XGB, and LR. The model archived AUC-ROC curve score is 0.998, and the PR-Curve is about 0.96, which is represented in the above Figure 14 It is very effective in real-time tasks. Because of its ensemble nature, which lowers bias and variation, it is useful for predicting mortality in the actual world. In soft voting, each classifier predicts a probability distribution over the possible output classes.

$$\hat{y} = arg \underset{c}{max} \sum_{i=1}^{T} w_i . p_{i,c}(x) \tag{7}$$

The variables $p_{i,c}(x)$, wi, and ŷ represent the expected probabilities of class c, weighted by classifier i, and the final predicted class, respectively.

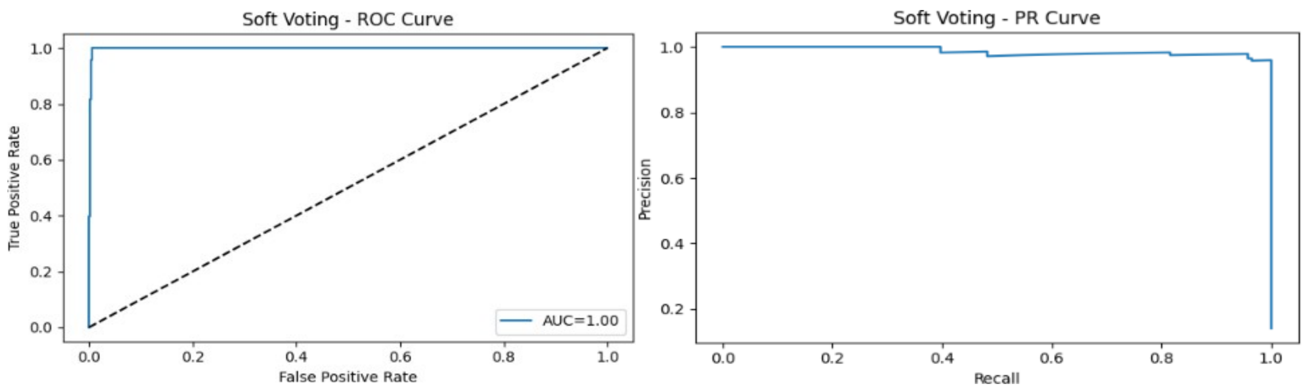


Fig. 14. Softvoting_ROC_Curve and PR_Curve for child mortality

L. Hard Voting

In the figure. Depicts hard voting for child mortality using AUC-ROC curve and PR-Curve. The ROC (0.9474) and PR curves show that hard voting, which makes predictions based on majority class labels, performs somewhat worse than soft voting. The above figure shows that this model uses a majority voting approach from the base classifiers. It chooses the class with a majority vote. It has been observed that the PR-AUC score is 0.90 and the ROC-AUC score is 0.94. In hard voting, each individual base classifier $h_i(x)$ provides a predicted class label, and the final predicted class is ŷ.

$$\hat{y} = mode\big(h_1(x), h_2(x), h_3(x), \dots \dots, h_n(x)\big) \tag{8}$$

Where 'n' is the total number of classifiers, and $h_i(x)$ denotes the class predicted by the i^{th} classifier.

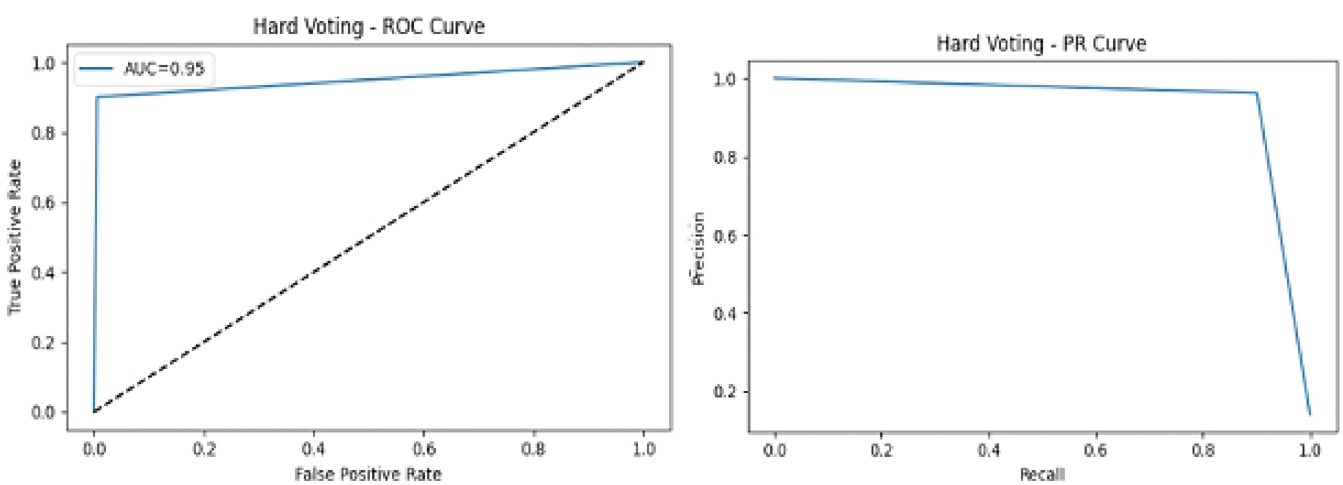


Fig. 15. Hard voting_ROC and PR_curve for child mortality prediction

M. Stacking

Stacking achieved outstanding results (ROC AUC 0.9999, PR AUC 0.9850) by combining numerous base classifiers and a Logistic Regression meta-classifier. It identifies both low- and high-level patterns, making it one of the most promising models for predicting child death. This model (Figure 16) learns the best combination of multiple models' outputs to predict child mortality. Trains the meta model on predictions from the base models.

A perfect accuracy score should be substantiated to rule out overfitting. we have applied the 10-fold cross-validation to all models. Average cross-validated metrics have been included in TABLE II. The confusion matrix for the top-performing models (AdaBoost, Stacking, and XGBoost) is shown in Figure 17.

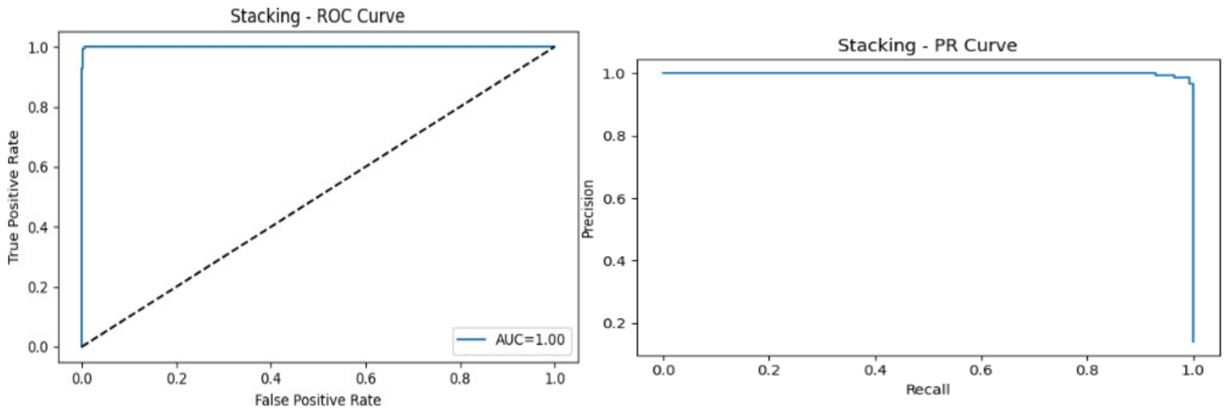
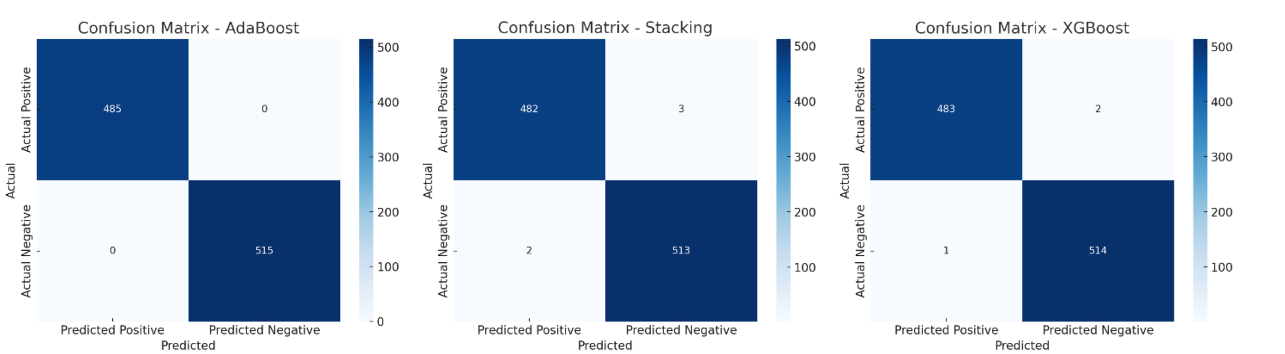


Fig. 16. Stacking AUC_ROC and PR_Curve for child mortality

TABLE II. AVERAGE 10-FOLD CROSS-VALIDATION OF DIFFERENT MACHINE LEARNING MODELS

S.No	Model	Precision	Recall	F1-Score	Accuracy	ROC AUC
1	Logistic Regression	0.7210	0.4278	0.5386	0.8942	0.9105
2	Decision Tree	0.9302	0.9181	0.9241	0.9773	0.9541
3	Random forest	0.9625	0.8603	0.9076	0.9745	0.9953
4	KNN	0.7012	0.5414	0.6117	0.9018	0.9056
5	SVM	0.8571	0.6623	0.7481	0.9355	0.9734
6	Naive Bayes	0.7701	0.4355	0.5570	0.9002	0.9359
7	Gradient Boosting	0.9962	0.9392	0.9668	0.9903	0.9998
8	XG-Boost	0.9837	0.9840	0.9838	0.9947	1.0000
9	Extra Trees	0.9502	0.7721	0.8507	0.9624	0.9912
10	AdaBoost	1.0000	1.0000	1.0000	1.0000	1.0000
11	Soft Voting	0.9664	0.9527	0.9595	0.9885	0.9980
12	Hard Voting	0.9543	0.8960	0.9241	0.9783	0.9458
13	Stacking	0.9758	0.9907	0.9832	0.9955	0.9999



(a) (b) (c)

Fig. 17. Confusion matrices (a) AdaBoost (b) Stacking (c) XGBoost.

N. Features most significantly influenced child mortality prediction

Figure 18 shows that most important features are used for child mortality prediction and would like to know which attribute is significantly influence when XGBoost algorithm applied on the dataset and how it is going to handle the class imbalance problems. The result indicates that XGB provides evidence of accuracy is 99.6 % and an F1-score is 98%.XGB is one of the classifiers that influences more for child mortality prediction. Additionally, the model's internal feature attribution mechanisms reinforce its selection for both predictive performance and interpretability by offering insightful information about maternal, socioeconomic, and healthcare-related variables have the greatest impact on mortality outcomes like wealth_index_poor and vaccination_status give much attention to child mortality prediction.

Limitations:

Potential overfitting due to limited sample diversity. The dataset's representativeness is limited because it is sourced from a single national survey. The absence of external validation, such as a holdout test set or external dataset, is a concern.

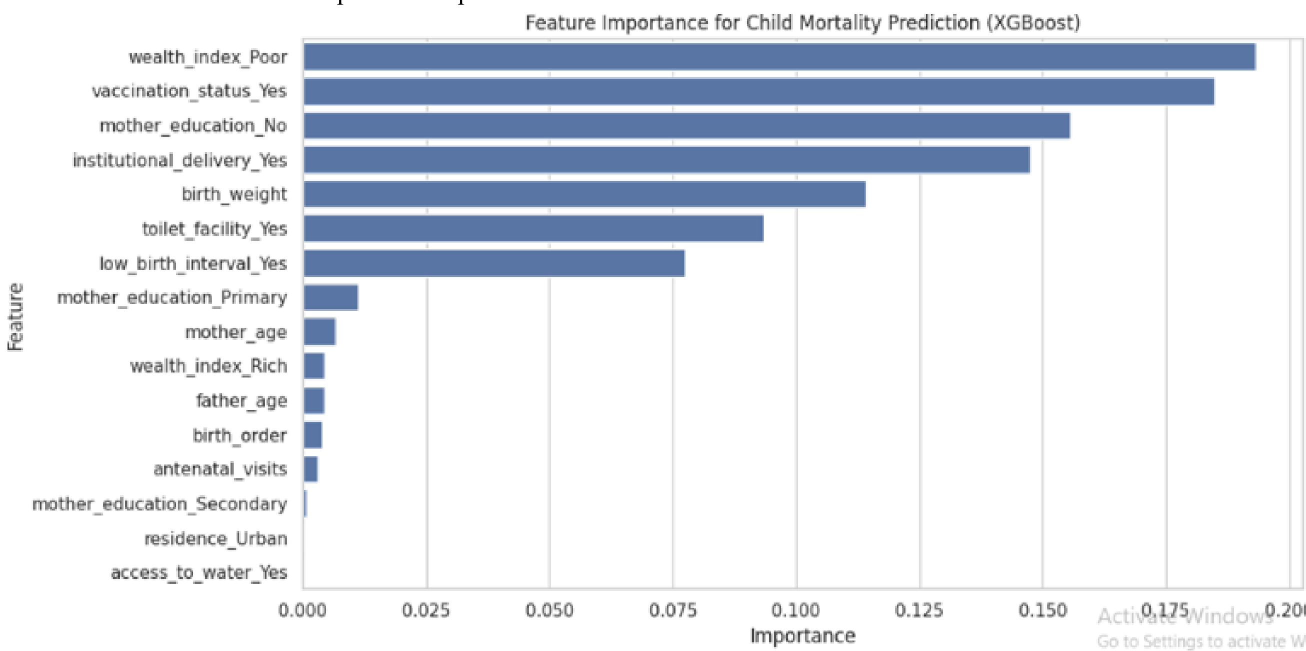


Fig. 18. Feature importance for child mortality prediction

V. Conclusion

This study report employed various machine learning techniques for predicting child mortality.. We considered 14 key demographic and health-related features, along with 13 ML classifiers used and evaluated the performance metrics to know which classifier performs well for child mortality. The performance metrics used. The experimental work reveals that AdaBoost is the top performer and obtained 100%. It means AdaBoost is the best classifier as comparisons to other classifiers. Apart from this, we also found XGBoost (Accuracy: 99.6%, F1-score: 0.9858) and Stacking (Accuracy: 99.6%, F1-score: 0.9859), showing the effectiveness of ensemble methods. A high-performing, deployable, and interpretable machine learning model for early child mortality prediction is provided by this study and can be included in national health systems. To guarantee scalable deployment, future enhancements can include adding temporal features, incorporation of explainable AI methods real-time data pipelines, and assessing the model's generalizability across regional datasets.

Conflict of Interest: All authors have stated that they have no competing interests.

Ethical Statement: The dataset used in this work is anonymized and publicly available, containing no personally identifiable information. Therefore, ethical approval or consent was not applicable.

References

[1] C. Di Novi and H. Thakare, "Inequality of Opportunity in Accessing Maternal and Newborn Healthcare Services: Evidence from the Bangladesh Demographic and Health Survey," *Social Indicators Research*, vol. 164, no. 3, pp. 1505–1529, Sep. 2022, doi: 10.1007/s11205-022-03004-6.

[2] F. Yu, Z. Yan, R. Pu, S. Tang, B. Ghose, and R. Huang, "Do Mothers with Lower Socioeconomic Status Contribute to the Rate of All-Cause Child Mortality in Kazakhstan?," *BioMed Research International*, vol. 2018, pp. 1–8, Jan. 2018, doi: 10.1155/2018/3629109.

[3] "End preventable deaths of newborns and children under 5 years of age," Mar. 14, 2024. https://www.who.int/data/gho/data/themes/topics/sdg-target-3_2-newborn-and-child-mortality

[4] E. K. Ekwuazi, C. O. Chigbu, and N. C. Ngene, "Reducing maternal mortality in low- and middle-income countries," *Case Reports in Women S Health*, vol. 39, p. e00542, Sep. 2023, doi: 10.1016/j.crwh.2023.e00542.

[5] P. Pandey, S. Shukla, N. K. Singh, and M. Kumar, "Predicting Child Mortality Determinants in Uttar Pradesh Using Machine Learning: Insights from the National Family and Health Survey (2019-21)," *Clinical Epidemiology and Global Health*, p. 101949, Feb. 2025, doi: 10.1016/j.cegh.2025.101949.

[6] F. Iqbal, M. I. Satti, A. Irshad, and M. A. Shah, "Predictive analytics in smart healthcare for child mortality prediction using a machine learning approach," *Open Life Sciences*, vol. 18, no. 1, 2023, doi: 10.1515/biol-2022-0609.

[7] A. W. Demsash, "Using best performance machine learning algorithm to predict child death before celebrating their fifth birthday," *Informatics in Medicine Unlocked*, vol. 40, p. 101298, 2023, doi: 10.1016/j.imu.2023.101298.

[8] C. Chivardi *et al.*, "Understanding the social determinants of child mortality in Latin America over the last two decades: a machine learning approach," *Scientific Reports*, vol. 13, no. 1, 2023, doi: 10.1038/s41598-023-47994-w.

[9] O. Samuel, T. Zewotir, and D. North, "Application of machine learning methods for predicting under-five mortality: analysis of Nigerian demographic health survey 2018 dataset," *BMC Medical Informatics and Decision Making*, vol. 24, no. 1, 2024, doi: 10.1186/s12911-024-02476-5.

[10] L. P. Garcia, I. J. C. Schneider, C. De Oliveira, E. Traebert, and J. Traebert, "What is the impact of national public expenditure and its allocation on neonatal and child mortality? A machine learning analysis," *BMC Public Health*, vol. 23, no. 1, 2023, doi: 10.1186/s12889-023-15683-y.

[11] S. Holcroft, I. Karangwa, F. Little, J. Behoor, and O. Bazirete, "Predictive modelling of postpartum haemorrhage using early risk factors: A comparative analysis of statistical and machine learning models," *International Journal of Environmental Research and Public Health*, vol. 21, no. 5, p. 600, 2024, doi: 10.3390/ijerph21050600.

[12] R. K. Saroj, P. K. Yadav, R. Singh, and Obvious. N. Chilyabanyama, "Machine Learning Algorithms for understanding the determinants of under-five Mortality," *BioData Mining*, vol. 15, no. 1, 2022, doi: 10.1186/s13040-022-00308-8.

[13] A. Satty, G. S. M. Khamis, Z. M. Mohammed, A. F. Mahmoud, F. A. Abdalla, M. Salih, ... & E. A. Gumma, "Statistical insights into machine learning models for predicting under-five mortality: An analysis from multiple indicator cluster survey (MICS)," *IEEE Access*, p. 1, 2025, doi: 10.1109/access.2025.3549097.

[14] S. Naznin, M. J. Uddin, I. Ahmad, and A. Kabir, "Analyzing and forecasting under-5 mortality trends in Bangladesh using machine learning techniques," *PLoS ONE*, vol. 20, no. 2, p. e0317715, 2025, doi: 10.1371/journal.pone.0317715.

[15] S. Das, L., D. Brals, B. Boczek, S. T. Hasan, P. Massara, ... & W. Voskuijl, "Development of machine learning models predicting mortality using routinely collected observational health data from 0-59 months old children admitted to an intensive care unit in Bangladesh: critical role of biochemistry and haematology data," *BMJ Paediatrics Open*, vol. 8, no. 1, p. e002365, 2024, doi: 10.1136/bmjpo-2023-002365.

[16] S. Adithya, T. S. Rezinald, A. V. A. Mary, J. Refonaa, S. L. J. Shabu, and P. Jeyanthi, "Child Mortality Prediction using Machine Learning," *2022 6th International Conference on Trends in Electronics and Informatics (ICOEI)*, vol. 5, pp. 1253–1256, 2022, doi: 10.1109/icoei53556.2022.9776827.

[17] C. Ashwini, S. R. Bose, M. S. D. Padmavathy, C. Raj, and J. C. Ajay, "Predicting child mortality with diverse regression algorithms using a machine learning approach," in *Advances in computer and electrical engineering book series*, 2024, pp. 329–352. doi: 10.4018/979-8-3693-3739-4.ch017.

[18] J. Prithula *et al.*, "Improved pediatric ICU mortality prediction for respiratory diseases: machine learning and data subdivision insights," *Respiratory Research*, vol. 25, no. 1, May 2024, doi: 10.1186/s12931-024-02753-x.

[19] M. H. Abdullah and Y. T. Mustafa, "Machine Learning and SHAP-Based analysis of deforestation and forest degradation dynamics along the Iraq–Turkey border," *Earth*, vol. 6, no. 2, p. 49, Jun. 2025, doi: 10.3390/earth6020049.

[20] A. A. Abdullah, M. M. Hassan, and Y. T. Mustafa, "Leveraging Bayesian deep learning and ensemble methods for uncertainty quantification in image classification: A ranking-based approach," *Heliyon*, vol. 10, no. 2, p. e24188, Jan. 2024, doi: 10.1016/j.heliyon.2024.e24188.

[21] A. A. Abdullah, M. M. Hassan and Y. T. Mustafa, "A Review on Bayesian Deep Learning in Healthcare: Applications and Challenges," in *IEEE Access*, vol. 10, pp. 36538-36562, 2022, doi: 10.1109/ACCESS.2022.3163384.

[22] M. S. Rao, S. K. UmaMaheswaran, N. C. Sattaru, K. H. Abdullah, U. K. Pandey and L. Biban, "A Critical Understanding of Integrated Artificial Intelligence Techniques for the Healthcare Prediction System," 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2022, pp. 728-731. https://doi: 10.1109/ICACITE53722.2022.9823678.

[23] B. A. Sullivan *et al.*, "Comparing machine learning techniques for neonatal mortality prediction: insights from a modeling competition," *Pediatric Research*, Dec. 2024, doi: 10.1038/s41390-024-03773-5.

[24] J. Lee, J. Cai, F. Li, and Z. A. Vesoulis, "Predicting mortality risk for preterm infants using random forest," *Scientific Reports*, vol. 11, no. 1, Mar. 2021, doi: 10.1038/s41598-021-86748-4.

[25] C. Mangold, S. Zoretic, K. Thallapureddy, A. Moreira, K. Chorath, and A. Moreira, "Machine Learning Models for Predicting Neonatal Mortality: A Systematic review," *Neonatology*, vol. 118, no. 4, pp. 394–405, Jan. 2021, doi: 10.1159/000516891.

[26] K. . Pal, J. . K Pal, S. R. Trivedi, N. Sharma, P. . P Patavardhan, and S. Goswami, "Predictive Modeling of Neonatal Mortality Trends Using Machine Learning: A Cross-sectional Data Science Approach", *J Neonatal Surg*, vol. 14, no. 22S, pp. 1033–1039, May 2025.

[27] A. Sheikhtaheri, M. R. Zarkesh, R. Moradi, and F. Kermani, "Prediction of neonatal deaths in NICUs: development and validation of machine learning models," *BMC Medical Informatics and Decision Making*, vol. 21, no. 1, Apr. 2021, doi: 10.1186/s12911-021-01497-8.

[28] S. K. Dey *et al.*, "Analyzing infant Cry to detect birth asphyxia using a hybrid CNN and feature extraction approach," *Neuroscience Informatics*, p. 100193, Feb. 2025, doi: 10.1016/j.neuri.2025.100193.

[29] M. S. Rao, Ch. Sekhar, and D. Bhattacharyya, "Comparative analysis of machine learning models on loan risk analysis," in *Advances in intelligent systems and computing*, 2021, pp. 81–90. https://doi: 10.1007/978-981-15-9516-5_7.

[30] A. F. M. Batista, C. S. G. Diniz, E. A. Bonilha, I. Kawachi, and A. D. P. C. Filho, "Neonatal mortality prediction with routinely collected data: a machine learning approach," *BMC Pediatrics*, vol. 21, no. 1, Jul. 2021, doi: 10.1186/s12887-021-02788-9.

[31] E. Keles and U. Bagci, "The past, current, and future of neonatal intensive care units with artificial intelligence: a systematic review," *Npj Digital Medicine*, vol. 6, no. 1, Nov. 2023, doi: 10.1038/s41746-023-00941-5.

[32] F. Juraev, S. El-Sappagh, E. Abdukhamidov, F. Ali, and T. Abuhmed, "Multilayer dynamic ensemble model for intensive care unit mortality prediction of neonate patients," *Journal of Biomedical Informatics*, vol. 135, p. 104216, Oct. 2022, doi: 10.1016/j.jbi.2022.104216.

[33] T. Włodarczyk *et al.*, "Machine Learning Methods for Preterm Birth Prediction: A review," *Electronics*, vol. 10, no. 5, p. 586, Mar. 2021, doi: 10.3390/electronics10050586.

[34] A. Rahman and Md. H. Rahman, "Explore the factors related to the death of offspring under age five and appraise the hazard of child mortality using machine learning techniques in Bangladesh," *BMC Public Health*, vol. 25, no. 1, Jan. 2025, doi: 10.1186/s12889-025-21460-w.

[35] N. Rai, N. Kaushik, D. Kumar, C. Raj, and A. Ali, "Mortality prediction of COVID-19 patients using soft voting classifier," *International Journal of Cognitive Computing in Engineering*, vol. 3, pp. 172–179, Jun. 2022, doi: 10.1016/j.ijcce.2022.09.001.