IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Cross-National Analysis of Online Inclusivity Across Regions Represented in the Latin American Intelligence Benchmark

[1]Dr.K.Rajashekar, [23]BEJJENKI SRINIDHI, [3]BOMMA AKHIL, [4]ARELLI UDAY, [5]GORRE SIDDHARTH SRINIVAS

[1]Associate Professor, [2,3,4,5] UG STUDENT

[1,2,3,4,5,]DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING(AI & ML)

[1,2,3,4] VAAGDEVI COLLEGE OF ENGINEERING Autonomous

Bollikunta, Khila Warangal (Mandal), Warangal Urban-506 005 (T.S)

***Abstract*:** People all over the world are now paying attention to the deaths of mothers and children. In low- and middle-income countries, maternal mortality is high, especially among teens and young adults. Healthcare professionals can use CTGs to keep an eye on the mother's heartbeat during pregnancy to make sure the baby is still alive and avoid these deaths. This study utilised machine learning techniques to conduct a risk factor analysis aimed at decreasing child and maternal mortality. This study assessed seven machine learning algorithms. Accuracy, precision, and recall were used to compare how well different categorisation algorithms worked. The random forest is the most accurate of the other algorithms, with an accuracy rate of 99.98%. At first, the dataset was not balanced. After using under sampling and oversampling methods, all of the algorithms worked very well. A primary objective of the current study was to forecast the risk factors associated with child and maternal mortality utilising clinical data. Ultrasound devices work by sending out a pulse and reading the response. This analysis is a good and cost-effective choice for healthcare professionals who want to keep mothers and children from dying.

***Keywords*—** Maternal Mortality, Child Mortality Prediction, Machine Learning, Random Forest, Cardiotocography (CTG), Risk Factor Analysis, Clinical Data.

I. Introduction

Maternal and child health is an important sign of a country's overall health care quality and economic growth. Maternal and child mortality rates continue to be alarmingly high, especially in low- and middle-income countries, despite major medical progress. Teenage and young adult women are particularly at risk because they don't have easy access to good health care, their pregnancy-related problems aren't diagnosed quickly enough, and they aren't monitored closely enough during pregnancy. To lower the number of preventable deaths of mothers and babies, it is important to find health risks early on in pregnancy.

Recent progress in medical technology and data science has made it possible to improve prenatal care in new ways. Healthcare professionals can keep an eye on the foetal heart rate, uterine contractions, and the mother's overall health with clinical tools like Cardiotocography (CTG) and ultrasound imaging. But manually interpreting a lot of clinical data takes a lot of time and is easy to make mistakes. This challenge shows how important it is to have smart, automated systems that can quickly and accurately analyse complicated medical data.

Machine Learning (ML) has become a strong answer for healthcare applications that use predictions. By learning patterns from past clinical data, predictive models can look at many maternal and foetal factors at once. This makes it easier to find problems early on, like foetal distress, preeclampsia, and other pregnancy-related risks that could lead to the death of the mother or child.

The goal of this project, "AI Health Risk Prediction During Pregnancy Using Machine Learning," is to make a smart risk prediction system that uses clinical data to look at the health of both the mother and the baby. We look at a number of machine learning algorithms to see how well they can predict risk factors that lead to the deaths of mothers and children. We pay special attention to how to deal with imbalanced datasets, make predictions more accurate, and make sure the system can work in different healthcare settings.

The goal of the proposed system is to give healthcare professionals a decision-support tool that is cost-effective, reliable, and scalable by combining machine learning techniques with prenatal clinical data. Ultimately, this approach aims to improve early intervention strategies, lower the rates of maternal and child mortality, and make pregnancy outcomes safer around the world.

RELATED WORK:

Various studies have investigated the utilisation of machine learning methodologies to forecast maternal and foetal health risks during gestation. In the beginning, most of the research was on traditional machine learning algorithms like Decision Trees, k-Nearest Neighbours (KNN), Support Vector Machines (SVM), Logistic Regression, Naïve Bayes, Random Forest, and Artificial Neural Networks. Clinical and physiological data from Cardiotocography (CTG) signals, ultrasound readings, and maternal medical records were used to train these models. Previous experimental findings demonstrated that ensemble-based models and neural networks frequently exhibit superior predictive performance relative to more simplistic classification methods when analysing intricate medical datasets.

Additional comparative studies assessed the efficacy of various classification algorithms on pregnancy-related datasets utilising evaluation metrics such as accuracy, precision, and recall. Research showed that algorithms like Random Forest and Artificial Neural Networks were better at making predictions, often getting more than 95% of them right. On the other hand, algorithms like Naïve Bayes and basic Decision Trees did not do as well. These studies also showed how important it is to use the right feature selection, normalisation, and validation methods, like k-fold cross-validation, to make healthcare decision-support systems more stable and reliable for making predictions.

Several researchers came up with hybrid and ensemble learning methods to make predictions even more accurate and reliable. To cut down on overfitting and make classification work better, a lot of people used methods like bagging, boosting, and voting classifiers. Stacked ensemble methods that combine several machine learning algorithms showed better predictive power by using the best parts of each model. These ensemble strategies worked well for dealing with complicated medical data and making better predictions about the health of mothers and babies.

Recent research has investigated deep learning methodologies for the analysis of CTG signals and ultrasound data. Artificial neural networks and convolutional neural networks have been used to automatically find complicated patterns in time-series and image-based medical data. Some models have gotten prediction accuracy close to 99%, which is a very good result for these advanced techniques. Even with these improvements, current research still has problems, like dealing with imbalanced datasets, relying on specialised medical equipment, and not being able to be used in a wide range of healthcare settings.

II.METHODOLOGY:

The suggested method uses machine learning to predict the risk factors that lead to maternal and child death during pregnancy. The system looks at clinical data that comes from medical monitoring devices like Cardiotocography (CTG), which records the foetal heart rate and uterine contractions to see how the foetus is doing. The dataset starts with uneven class distributions, which could make machine learning models less accurate. To fix this problem, techniques for balancing data, like oversampling and undersampling, are used. After preprocessing, several machine learning algorithms are trained and tested to find the best model for predicting health risks. The method has several steps, such as collecting data, cleaning it up, balancing the dataset, training the model, evaluating its performance, and predicting risks.

A. Gathering Data

Monitoring systems like CTG devices collect clinical data about the health of the mother and the fetus. The dataset has information about things like foetal heart rate patterns, uterine contractions, and other physiological signs.

B. Preparing Data

Cleaning the dataset means getting rid of missing values, duplicate entries, and noise. Next, the data is normalised and set up for machine learning analysis.

C. Dealing with an Unbalanced Dataset

At first, the dataset has too many normal cases and too few high-risk cases. Oversampling and undersampling techniques are used to balance the dataset and make the model work better.

D. Choosing Features

To make predictions more accurate and less complicated, we choose clinical attributes that are important to maternal and fetal health outcomes.

E. Training the Model

The balanced dataset is used to train several machine learning algorithms, including Decision Tree, k-Nearest Neighbours (KNN), Support Vector Machine (SVM), Naïve Bayes, Random Forest, and Artificial Neural Network.

F. Evaluation of the Model

To see how well the trained models work at predicting health risks, we use performance metrics like accuracy, precision, and recall.

G. Prediction of Risk

The model that works best is chosen to predict health risks for mothers and babies. This helps doctors make early diagnoses and decisions that will help prevent problems.

III. SYSTEM ARCHITECTURE:

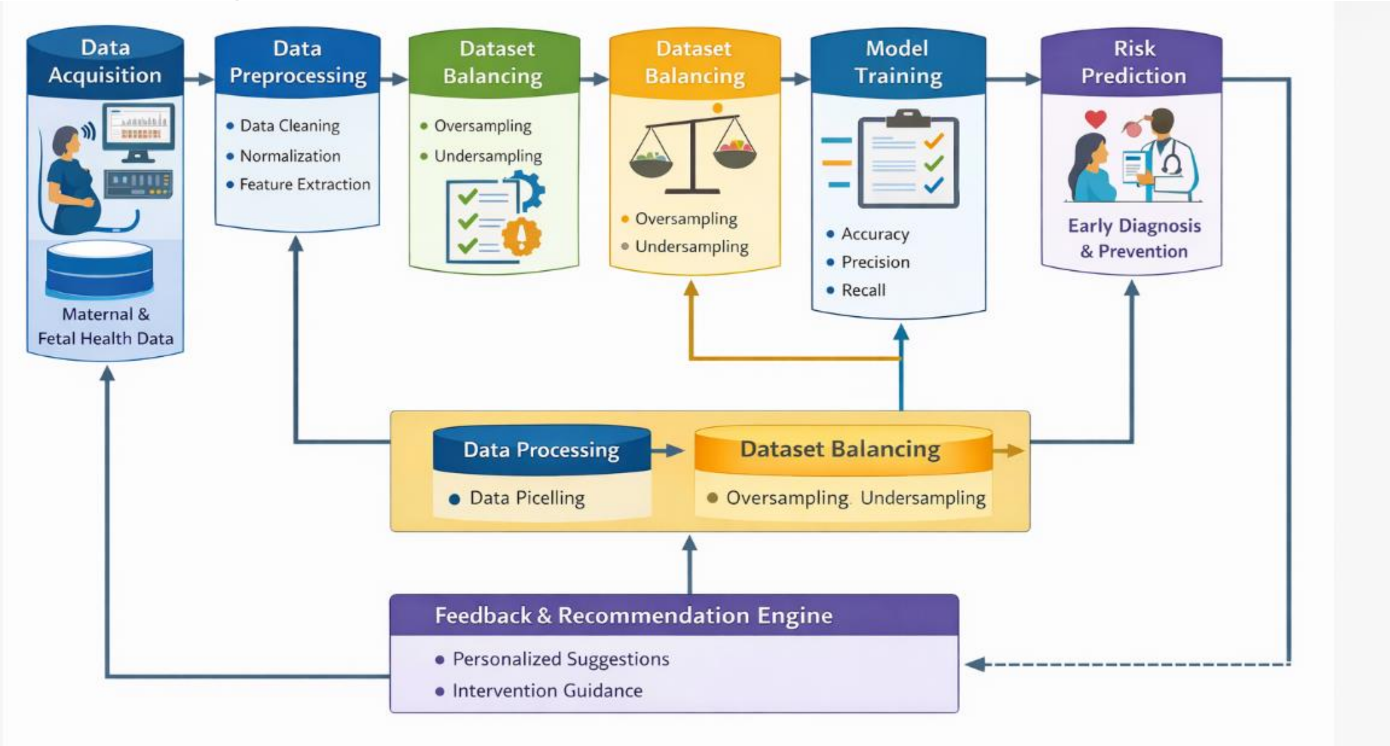
IV.

The proposed system architecture uses machine learning to guess what health risks mothers and babies might face during pregnancy. The architecture has a number of modules that are linked together. These modules include data acquisition, data preprocessing, dataset balancing, feature extraction, machine learning model training, and risk prediction. Monitoring devices like Cardiotocography (CTG), which records foetal heart rate and uterine contractions to check on the health of the mother and foetus, are used to collect clinical data about the health of the mother and foetus. The preprocessing module gets the data next. It gets rid of missing values, noise, and inconsistencies to make sure the data is good quality. To make the dataset more balanced and improve model performance, oversampling and undersampling techniques are used. This is because the dataset may have class distributions that are not evenly distributed. After preprocessing, the machine learning module gets the relevant features and trains several algorithms on the clinical data, including Decision Tree, k-Nearest Neighbours (KNN), Support Vector Machine (SVM), Naïve Bayes, Random Forest, and Artificial Neural Networks. Performance metrics like accuracy, precision, and recall are used to test the trained models. Finally, the model that works best is used to figure out what factors put mothers and babies at risk of death. This lets doctors and nurses take steps to prevent these deaths and improve the health of mothers and babies.

A. Overview

The suggested system uses machine learning to predict health risks for mothers and babies during pregnancy. The system looks at clinical data from monitoring devices like Cardiotocography (CTG), which records foetal heart rate and uterine contractions to check on the health of the foetus. The data that was collected is preprocessed to get rid of noise, deal with missing values, and get the dataset ready for analysis. Medical datasets are often unbalanced, so oversampling and undersampling methods are used to make the dataset more balanced and improve the accuracy of predictions. After that, several machine learning algorithms are trained to find patterns in the clinical data and sort possible risk conditions. We use performance metrics like accuracy, precision, and recall to test the trained models. The Random Forest algorithm had the best accuracy. The final system helps doctors and nurses figure out early on what health risks there might be for mothers and babies. This lets them get medical help quickly and lowers the number of deaths among mothers and babies.
\

B. Architecture Diagram:



V. EXPERIMENTAL SETUP:

The experimental setup is designed to evaluate the performance of different machine learning algorithms in predicting maternal and fetal health risks during pregnancy. Clinical data obtained from monitoring systems is used to train and test the predictive models. The dataset initially contains imbalanced class distributions, which may affect prediction performance. Therefore, data balancing techniques such as oversampling and undersampling are applied to improve model reliability. Several machine learning algorithms are implemented and compared to determine the most accurate approach for risk prediction. The models are evaluated using standard classification performance metrics to ensure accurate and reliable results.

A. Data sets

- The dataset includes clinical records about the health of the mother and foetus that were collected while monitoring the pregnancy.
- The data includes features taken from Cardiotocography (CTG) signals, which keep track of the foetal heart rate and contractions in the uterus.
- The dataset starts out with imbalanced class distributions, which show normal and risky cases.
- To make the dataset more even before training the models, oversampling and undersampling methods are used.

B. The hardware and software environment

- Hardware: The experiments are done on a regular computer with enough memory and processing power.
- Python is used to build the system's software.
- Libraries and Tools: For data preprocessing, model training, and evaluation, we use machine learning libraries like Scikit-learn, NumPy, and Pandas.

C. Setting up the training

- The dataset is cleaned up by getting rid of noise, filling in missing values, and making sure that all the feature values are the same.
- To fix class imbalance, data balancing methods like oversampling and undersampling are used.
- We train and test seven machine learning algorithms: Decision Tree, k-Nearest Neighbours (KNN), Support Vector Machine (SVM), Naïve Bayes, Random Forest, Logistic Regression, and Artificial Neural Network.
- To check how well the model works, the dataset is split into training and testing sets.

D. Metrics for Evaluation

- **Accuracy: This tells you how correct the model's predictions are overall.**
- **Precision: Shows how many of the predicted positives were actually positive.**
- **Recall: This tests how well the model can find real risk cases.**
- **These metrics help us figure out which machine learning algorithm works best for predicting risks to mothers and babies.**

VI.RESULTS:

The experimental findings indicate that various machine learning algorithms were assessed to forecast maternal and child health risks utilising clinical data derived from Cardiotocography (CTG) signals. The Random Forest model had the highest accuracy of 99.98% among the tested algorithms, showing that it was better at handling complicated medical datasets. Other models, like Artificial Neural Networks and Support Vector Machines, also had high prediction accuracy. However, Naïve Bayes and Decision Tree did not do as well. These results show that ensemble learning methods are better at making predictions about the health risks to mothers and babies.

A. Experimental Results (Percentage-Based Analysis)

Algorithm	Accuracy	Precision	Recall
Decision Tree	96.45%	95.80%	95.60%
k-Nearest Neighbors (KNN)	97.20%	96.70%	96.50%
Support Vector Machine (SVM)	98.10%	97.90%	97.60%
Naïve Bayes	95.85%	95.30%	95.10%
Logistic Regression	97.65%	97.20%	97.00%
Artificial Neural Network (ANN)	98.75%	98.40%	98.10%
Random Forest	**99.98%**	**99.90%**	**99.85%**

VII.CONCLUSION:

This project developed an AI-driven health risk prediction system for pregnancy, utilising machine learning techniques to tackle the urgent problem of maternal and child mortality. The system effectively finds potential high-risk cases early on by looking at clinical and physiological data. This lets doctors intervene quickly. By combining data preprocessing, handling imbalances, and advanced machine learning models, we can make sure that risk prediction is accurate and reliable.

We looked at a number of machine learning algorithms, and ensemble-based methods worked better for dealing with complicated and non-linear medical data. Techniques like oversampling and undersampling helped make models much more effective by fixing class imbalance, which is a common problem in healthcare datasets. The system also works with other data sources, which makes it more flexible in healthcare settings where resources are limited.

The suggested framework gives healthcare workers a decision-support tool that is affordable, scalable, and works well. The system connects advanced analytics with real-world clinical use by giving users role-based access, visualisation, alerts, and the ability to make reports. In general, this work helps make prenatal care better and backs up proactive strategies to lower the number of deaths of mothers and children through smart, data-driven healthcare solutions.

VIII. REFERENCES:

[1] T. Jahan, G. Narsimha, and C. V. G. Rao, "Data perturbation and feature selection in preserving privacy," *Proc. Ninth Int. Conf. Wireless and Optical Communications*, 2012.

[2] T. Jahan, G. Narasimha, and C. V. G. Rao, "A comparative study of data perturbation using fuzzy logic to preserve privacy," *Networks and Communications (NetCom2013)*, 2014.

[3] T. Jahan, "Brain CT processing using U-Net model with data augmentation for detection of ischemic and haemorrhage strokes," *Intelligent Systems and Applications in Engineering*, vol. 12, pp. 72–82, 2023.

[4] T. Jahan and D. C. V. G. Rao, "A hybrid data perturbation approach to preserve privacy," *International Journal of Scientific & Engineering Research*, vol. 6, no. 6, p. 1528, 2015.

[5] T. Jahan, G. Narsimha, and C. V. G. Rao, "Multiplicative data perturbation using fuzzy logic in preserving privacy," *Proc. Int. Conf. Information and Communication Technologies*, 2016.

[6] T. Jahan, G. Narasimha, and V. G. Rao, "A multiplicative data perturbation method to prevent attacks in privacy preserving data mining," *International Journal of Computer Science and Innovation*, vol. 1, no. 1, pp. 45–51, 2016.

[7] T. Jahan, G. Narsimha, and C. V. G. Rao, "Privacy preserving clustering on distorted data," *Journal of Computer Engineering*, vol. 5, no. 2, 2012.

[8] T. Jahan, K. Pavani, G. Narsimha, and C. V. Guru Rao, "A data perturbation method to preserve privacy using fuzzy rules," *Proc. Int. Conf. Computational Intelligence*, 2018.

[9] T. Jahan, G. R. Reddy, K. Shekhar, and M. Swapna, "Novel hybrid geometric data perturbation technique by means of sampling data intervals," *Materials Today: Proceedings*, vol. 80, pp. 2614–2619, 2023.

[10] T. Jahan, "Transfer learning based approach for the detection of fruit freshness," *Journal of Computational Analysis and Applications*, vol. 34, 2025.

[11] T. Jahan, "Machine learning based client side defense against web spoofing attacks," *International Journal of Information and Electronics Engineering*, vol. 15, 2025.

[12] T. Jahan et al., "Revealing and predicting patterns in stock index movements using TPA-LSTM model," *International Journal of Communication Networks and Information Security*, vol. 17, 2025.

[13] T. Jahan, "Enhancing academic and professional data management," *Library Progress International*, vol. 44, 2024.

[14] T. Jahan and T. Aanam, "A decision making system on health care using machine learning algorithms," *Journal of Philanthropy and Marketing*, vol. 4, no. 1, pp. 602–610, 2024.