**Abstract**

Worldwide, maternal health during pregnancy is a major concern, especially in rural areas where risks are increased by a lack of medical experts and poor infrastructure. The lack of effective methods for predicting maternal health risks poses a significant challenge in maternal healthcare. Traditional approaches often fall short of accurately identifying and managing complications during pregnancy, leading to adverse outcomes. To close this gap, new methods must be developed that can reliably analyse medical data to predict risks to maternal health. This research focuses on investigating the utility of machine learning algorithms in predicting maternal health risks based on various medical parameters. By examining and contrasting several machine learning algorithms, this study aims to improve the efficacy and precision of maternal risk prediction. These algorithms are chosen for their ability to handle complex datasets and make accurate predictions based on maternal health parameters. By comparing the performance of these algorithms, this paper aims to identify the most effective approach for predicting maternal health risks. To train and test the predictive models, a carefully selected dataset with medical parameters related to maternal health has been taken into consideration in order to train and evaluate the predictive models. To guarantee the accuracy and dependability of the input data, this dataset—known as the Maternal Health Risk Dataset—is pre-processed utilizing sophisticated data cleaning and feature engineering techniques. This study uses an extensive set of data to find the most effective predictive model capable of accurately identifying maternal health risks, thereby reducing adverse outcomes during pregnancy and facilitating timely interventions when needed. The mentioned dataset comprises seven features: Age, Systolic BP, Diastolic BP, BS (Blood Glucose Levels), Body Temperature, Heart Rate, and Risk Level as target classes. Employing ensemble learning-based feature engineering, the study explores the efficacy of algorithms such as Random Forest, XG Boost, Support Vector Classifier (SVC), Decision Tree, and Logistic Multiclass. The achieved accuracies showcase the predictive performance of these algorithms, with Random Forest boasting an impressive accuracy of 94.26%.

**Introduction**

Pregnancy, a transformative phase in a woman's life, is often accompanied by a myriad of physiological changes and potential health risks. According to recent statistics, globally, approximately 295,000 women died due to pregnancy-related complications in 2017, with a significant portion of these deaths being preventable [1]. The World Health Organisation estimates that 810 women die every day from pregnancy- and childbirth-related preventable causes, with low- and middle-income countries accounting for 94% of these deaths. Teenagers and young adults under the age of twenty-one are particularly at risk [2]. Recognising the critical need for proactive management of maternal health, predictive modelling emerges as a promising approach to anticipate and mitigate pregnancy risks.

Individuals encounter considerable difficulties and incur preventable outcomes in the absence of risk-prediction pregnancy models. When these models are lacking, possible problems are not recognised early enough, which causes medical intervention to be insufficient or delayed [3]. This delay can exacerbate health issues for both the mother and the foetus, leading to adverse outcomes such as preterm birth, preeclampsia, and gestational diabetes [4]. Moreover, without predictive models, healthcare resources may not be allocated optimally, leading to inefficiencies and disparities in healthcare access [5]. Overall, the lack of risk prediction pregnancy models raises the rates of maternal death while also causing unnecessary pain and health issues for expectant mothers around the globe.

This research explores several cutting-edge machine learning techniques, including Random Forest with hyperparameter optimization, XGBoost, Support Vector Classification (SVC), Decision Tree, and Logistic Multiclass models.

The primary objective of this study is:

* To determine the most accurate and reliable Machine Learning methodology for predicting pregnancy risks, via rigorous comparison and analysis.
* The evaluation of multiple parameters, such as age, blood glucose levels, heart rate, risk level, systolic and diastolic blood pressure, and heart rate, in order to carry out a thorough comparison.

The objective of our study is to improve the sensitivity and accuracy of risk assessment during pregnancy by including several characteristics in our prediction models. This will allow for prompt intervention and individualized treatment.

The second section of the offered article discusses the commendable achievements, successes, and approaches used in similar investigations and research endeavors. In Section 3, the research technique and processes that were employed during the investigation are outlined in detail. Next, Section 4 provides a thorough summary of the work that has been done, including particulars, results from the experiments, and assessment criteria. Section 5, which concludes with a summary of the research findings and a discussion of potential directions for future research within the same theme subject, is the culmination of the completed work.

**Related Work**

MD Assaduzzaman et al. [6] predicted maternal health risk factors using a variety of machine learning techniques. The subsequent methodologies were employed: feature engineering, data cleansing, data collection and preparation, and the use of Cat Boost, Random Forest, XGB, Decision Tree, and Gradient Boost. The Random Forest algorithm did the best, scoring 90% accuracy.

Hursit Burak MUTLU et al. [7] calculated the risk to maternal health using machine learning techniques. In the study, six distinct machine learning techniques were applied. Six distinct features were obtained to estimate maternal health risk; and decision trees, light GBM Classifiers, catboosts, random forests, gradient boosting machines, and KNN classifiers were utilized to categorize the data in the dataset. With an accuracy of 89.16%, Decision Tree was shown to be the most effective technique. The k-nearest neighbours (KNN) approach yielded the lowest accuracy rate, at 68.47%.

According to Julio Jerison E. Macrohon et al. [8], predicting high-risk pregnancies is essential for the development of the unborn child and the health of the mother. The ability of supervised machine learning systems to forecast high-risk pregnancies was evaluated. The second-best algorithm for predicting high-risk pregnancies is Multilayer Perceptron. The third and fourth best algorithms are Support Vector Machine and Random Forest, respectively. In terms of accuracy, K-Nearest Neighbours and Naive Bayes did poorly. The decision tree algorithm scored 93.70% on the test, which was the highest. The accuracy of a semi-supervised method utilizing a Self-Training model was 97.01%. The work fills the knowledge gap regarding high-risk pregnancy prediction in cases of shaky or sparse data.

Ivana Maric et al. [9] makes reference to an early preeclampsia prediction model was created using machine learning techniques. Laboratory and clinical data from standard prenatal appointments was examined. Utilising a gradient boosting approach, the prediction model was constructed. The results were compared using the generalized linear model and logistic regression. A distinct model was created for preeclampsia with an early onset.

Lokesh Pawar et al. [10] suggested in their study that unsafe abortions, high blood pressure, bleeding, and early labour are the main causes of maternal mortality. Conventional machine learning techniques are used to predict the danger to maternal health. In order to function even in the worst, average, and best scenarios and deliver the robust performance by taking the performance of all scenarios into consideration, K-Fold Cross Validation has been proposed. With an accuracy of 70.21%, the suggested robust model proves to be the most effective robust model out of all of them. In this study, a conventional machine learning technique is utilized for training and testing, while the Gini index is used for feature selection.

|  |  |  |  |
| --- | --- | --- | --- |
| **Work** | **Dataset** | **No. Of Data** | **Accuracy** |
| [6] | Maternal Health Risk Dataset- Kaggle | 1014 observations (more details need to add for the dataset) | Random Forest accuracy : 90%, after data preprocessing. |
| [7] | Pima-Indian diabetes dataset used for risk factor analysis and comparison. | 203 observations  (more details need to add for the dataset) | Decision Tree method accuracy: 89.16%  K-nearest neighbours method accuracy: 68.47%  Logistic Model Tree method accuracy: highest accuracy. |
| [8] | -Limited collected data from the municipality of Daraga in Albay.  -Synthetic data generated using Synthetic Data Vault.  -Additional 900 rows generated for testing.  -Data balanced using Synthetic Minority Oversampling Technique (SMOTE) | original dataset had 90 rows; After data balancing using SMOTE, the dataset increased to 112 rows. | Decision Tree algorithm accuracy: 93.70% |
| [9] | data from Lucile Packard Children Hospital | -The initial cohort included 16,370 births.  -Preeclampsia was diagnosed in 561 pregnancies.  -The final cohort had data for 5,245 deliveries. | better accuracy in predicting early-onset preeclampsia compared to all preeclampsia. |
| [10] | Dataset is taken from UCI Machine Learning repository.  Data is collected from hospitals, maternal health cares, and community clinics in rural areas of Bangladesh. | 1014 instances with 6 features. | robust model has an accuracy of 70.21%. |

**Methodology**

This section sheds a light on research methodology, along with data collection, preprocessing, model development and its evaluation technique which were used for analysis and risk factor prediction.

**Data Set: Maternal Health Risk**

In this paper the maternal health risk factors dataset [28] was used for this study, which we got from the UCI machine learning public repository. The majority of the 1014 rows and 7 columns in this dataset reflect pregnant women who pose minimal health risks. Of the 1014 observations, 406 (40%) were classified as low risk, 336 (33.1%) as medium risk, and 272 (26.8%) as high-risk pregnant women. After that, balanced the dataset using categorical encoding and hyper tuned it, yielding a total of 1218 observations. The following provides a full description of each dataset attribute:

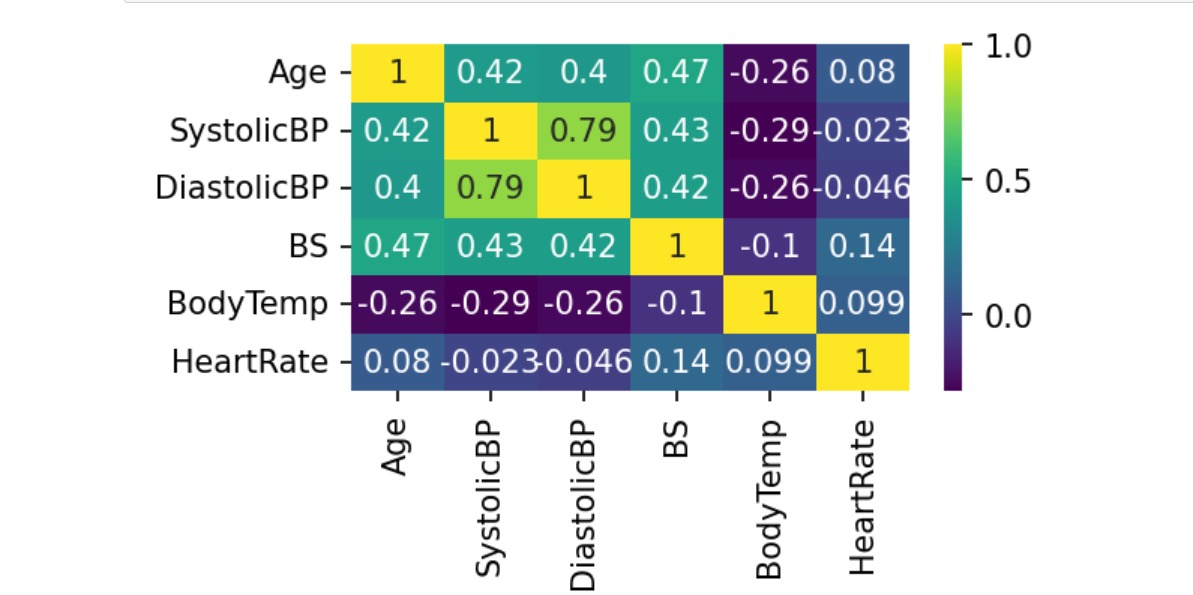
* **Age:** Represents the age of pregnant women in years.
* **Systolic BP:** Denotes the maximum blood pressure in millimeters of mercury, a crucial parameter during pregnancy.
* **Diastolic BP:** Indicates the lower blood pressure measurement in millimeters of mercury, another vital consideration during pregnancy.
* **BS:** Refers to the amount of glucose in the blood, measured in mmol/L.
* **Heart rate:** Represents the normal heart rate in beats per minute.
* **Risk Level:** Signifies the intensity level of risk prediction during pregnancy, dependent on the preceding attributes.

The dataset was split into two parts for machine learning purposes: 80% for training the model and 20% for testing its performance. This allows the model to learn from most of the data and then evaluate its effectiveness on unseen data, a standard practice in machine learning. The distribution of training and testing data across different categories is displayed in Table 1 below.

* **Table 1: Dataset Description**

|  |  |  |
| --- | --- | --- |
|  | Train | Test |
| High Risk Level | 324 | 82 |
| Mid Risk Level | 324 | 82 |
| Low Risk Level | 324 | 82 |

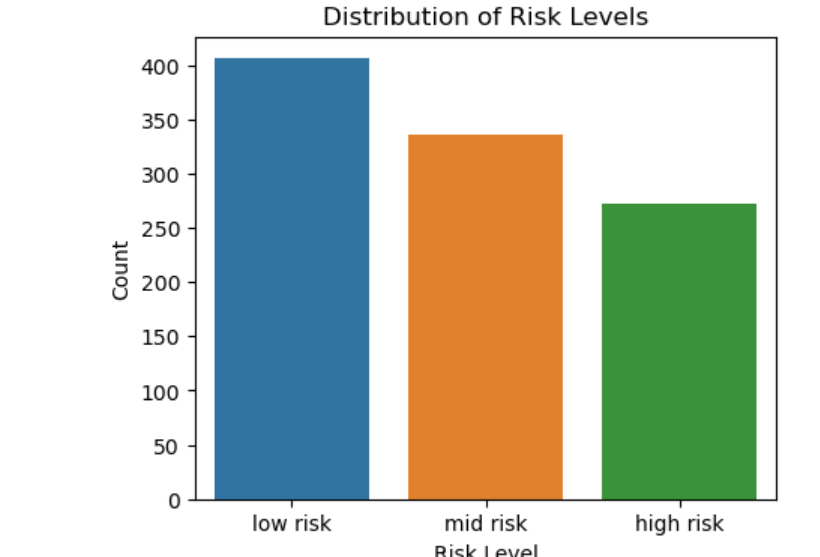
In this context, the target variable is the risk level, while the remaining features serve as predictor variables. The heat map for the dataset's correlation features is shown in Figure 1 below.



**Figure 1:** The heat map for correlation features of the dataset.

**Data Preprocessing:**

The initial phase of enhancing the predictive model involves ensuring that the data is meticulously prepared for analysis. The effectiveness of the model is substantially heightened when the data undergoes appropriate transformations. The dataset exhibited a commendable absence of missing values, reflecting a high degree of data completeness. But when we looked at the kinds of pregnancies in our dataset, we noticed that only 26.8% of them were considered high-risk. This imbalance might affect our analysis, so we need to be careful when drawing conclusions about risk factors. The figure below shows the distribution of the data frequencies in three categories: high, low, and medium.



**Figure 2:** Risk Level Distribution

**Categorical Encoding**

In the preprocessing phase, we conducted categorical encoding on the 'RiskLevel' variable, converting it into numerical values. This transformation is crucial for machine learning algorithms, which typically require numerical inputs for optimal performance. By ensuring compatibility and effectiveness in subsequent analyses, this encoding step facilitates seamless integration of the 'RiskLevel' variable into our machine learning models, contributing to improve their overall accuracy and predictive capabilities.

**Feature Selection**

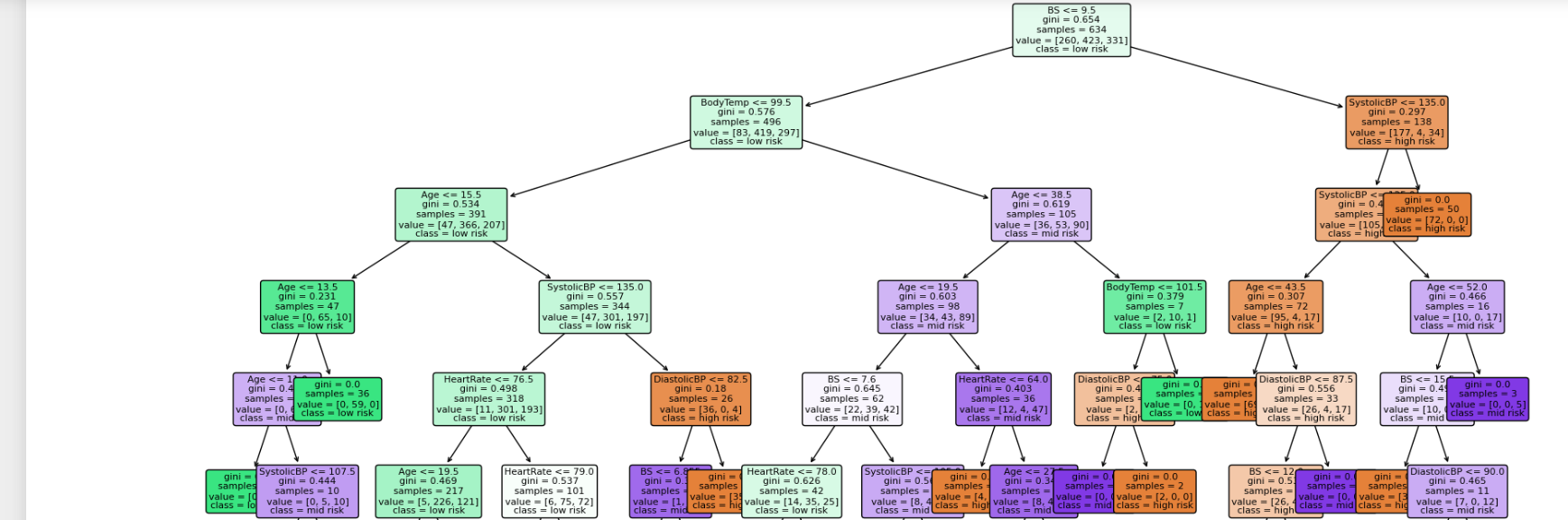
In the process of feature selection, a strategic decision was made to eliminate the 'DiastolicBP' column from our dataset. Feature selection is a crucial preprocessing technique in machine learning aimed at optimizing model performance. By carefully choosing relevant features and discarding less impactful ones, we can streamline our data, reducing complexity and potentially enhancing the predictive capabilities of our models. This approach is geared towards improving model efficiency and accuracy by focusing on the most informative variables, a fundamental practice in machine learning engineering

**Handling Class Imbalance**

Addressing class imbalance is pivotal for an unbiased model. The Synthetic Minority Over-Sampling Technique (SMOTE) was employed to rectify the class imbalance in the target variable 'RiskLevel'. SMOTE adeptly generates synthetic samples for the minority class, guaranteeing a balanced representation within the dataset.

**Proposed Methodology**

Several machine learning methods, including Random Forest, Support Vector Classification (SVC), and XGBoost, were initially considered. Following a thorough evaluation, the Random Forest Classifier emerged as the top performer among these methods. Figure below provides a detailed visualization of a decision tree within the Random Forest ensemble, featuring Gini impurity values at various nodes. The Gini impurity serves as a measure of the model's decision-making process, highlighting key points where feature values are evaluated for optimal classification.

****

**Figure 3 : Decision Tree with Gini Impurity in Random Forest**

**Random Forest Classifier:**

The Random Forest Classifier, our chosen predictive model, excels as an ensemble learning method by aggregating insights from multiple decision trees, thereby enhancing predictive accuracy and addressing overfitting concerns.

**Hyperparameter Tuning:**

Optimizing the Random Forest Classifier involved the implementation of Grid Search, an advanced hyperparameter optimization technique. This approach systematically explored various hyperparameter configurations, with a focus on key parameters like the number of estimators (trees), maximum tree depth, minimum samples for node splitting, and minimum samples for a leaf. The goal was to fine-tune these parameters for the optimal model performance.

**Model Training:**

Following hyperparameter tuning, the Random Forest Classifier underwent intensive training using the preprocessed dataset. The optimal hyperparameter configuration derived from the Grid Search process was applied, ensuring the model is finely tuned for superior performance. This well-trained classifier forms the cornerstone for subsequent evaluation and testing stages, showcasing its efficacy in handling real-world data with precision and reliability.

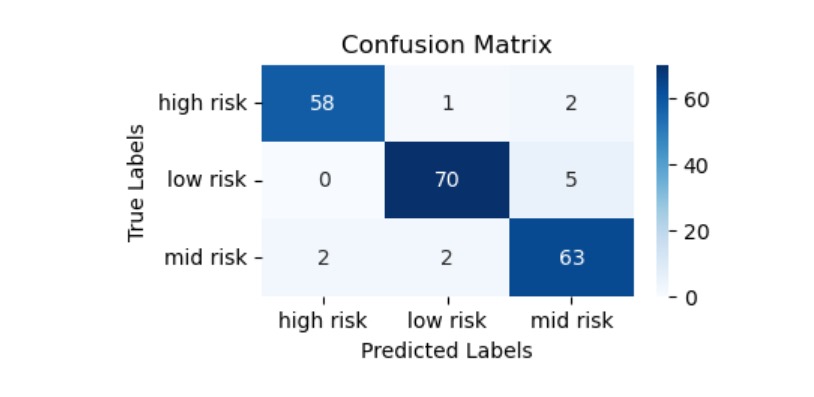
**Experimental Analysis**

In this section, the following analysis shows the experimental results, and provide insights of graphs into the implications of our analysis for the Maternal Health Risk. Experiment configuration is as follows:

* **Hardware:** AMD Ryzen 7, 1.90 GHz (16 GB RAM)
* **Software:** Google Colab CPU, T4 GPU
* **Libraries: Matplotlib,Seaborn,Scikit-learn**
* **Architectures used: Random Forest**
* **Dataset:** Maternal Health Risk

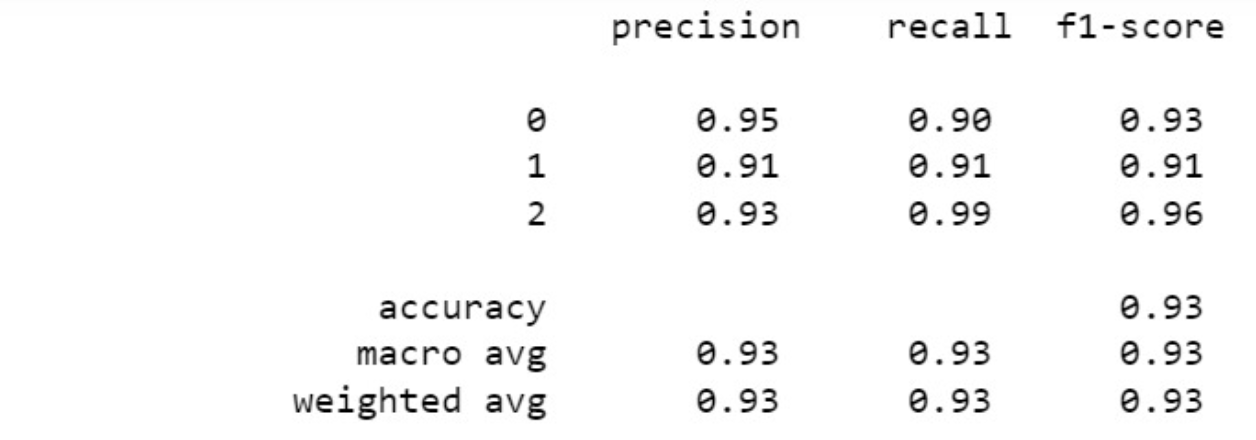
Three basic classification measures were used to evaluate the model's performance: accuracy, precision, and recall. Recall rated the model's capacity to identify positive instances, accuracy scored overall correctness, and precision assessed the significance of positive predictions.   
  
In order to improve the resilience of the model assessment and reduce the possibility of overfitting, a strict 5-fold cross-validation technique was employed throughout the hyperparameter adjustment stage. To do this, the dataset was divided into five subsets. The model was then trained on four of the subsets, and its performance was assessed on the fifth subset. To offer a thorough evaluation of the model's capabilities, the procedure was done five times, with the average results.

A thorough classification report and important metrics like accuracy, precision, and recall were included in the final reporting of the model's performance. A comprehensive evaluation viewpoint was provided by the classification report, which was crucial in providing detailed insights into the model's performance across several classes. Confusion matrices, which separate predictions into true positives, true negatives, false positives, and false negatives, were also used to further explain the model's performance. The confusion matrix, which provides a visual depiction of the Random Forest model's classification performance, is shown in the figure below. True positive, true negative, false positive, and false negative predictions are briefly summarized, offering important information about the model's accuracy and error rates.



**Figure 3:** Normalized Confusion matrix of Random Forest model

The accuracy percentages for the different machine learning algorithms used to forecast the hazards to maternal health are shown in the table below, along with the associated Precision, Recall, and F1 Score values. Notably, the model's impressive accuracy rate of 94.26% shows how accurate it is generally at predicting the consequences of maternal health. Confusion matrices were used in addition to these metrics to offer a detailed analysis of the model's performance, assisting in the discovery of regions in which the model performed particularly well or encountered difficulties in predicting particular outcomes. This all-encompassing strategy made sure that the algorithms and their efficacy in addressing maternal health risk prediction were thoroughly evaluated. The precision ,recall and f1 score for each risk category is displayed in the figure below.



**Figure 4: Comparative Analysis of Precision, F1-Score, and Recall for various Categories**

**Conclusion**

In the final analysis, our work has made an ambitious attempt to predict the health risks to mothers using a variety of machine learning methods. The findings of our study show that random forest with hyperparameterization is more effective than other models. A noteworthy achievement is the model's astounding accuracy percentage of 94.26%.  
  
Even though these findings are encouraging, it's important to recognize some limitations. The quality and availability of data, as well as regional differences in healthcare procedures, limited the study's breadth. Additionally, modifications to healthcare regulations and the evolution of diagnostic standards may have an impact on the model's performance.

There appears to be a bright future for additional developments in this area. Subsequent investigations may examine the incorporation of supplementary features or datasets to augment prediction capacities. Furthermore, the implementation of real-time data streams and ongoing model updates may enhance the responsiveness and dynamic nature of maternal health risk assessment.

As a result, our model's performance represents a major breakthrough in the area of specialized medical treatments for pregnant mothers. Application of advanced computational technologies could revolutionize prenatal care by enabling early identification and mitigation of maternal health concerns. The present study establishes the foundation for further endeavors aimed at enhancing and expanding the use of machine learning in maternal healthcare. These endeavors are anticipated to provide superior outcomes for prenatal care and heightened maternal welfare.

--------------------------

Sources-

[1] Maternal mortality – WHO

[2] World Health Organization. Maternal Mortality. Available online: https://www.who.int/en/news-room/fact-sheets/detail/ maternal-mortality

[3] A Semi-Supervised Machine Learning Approach in Predicting High-Risk Pregnancies in the Philippine

[4] Improving preeclampsia risk prediction by modeling pregnancy trajectories from routinely collected electronic medical record data- <https://www.nature.com/articles/s41746-022-00612-x>

[5] https://www.publichealth.columbia.edu/research/population-health-methods/risk-prediction

[6] Predicting Hypertensive Disorders in High-risk Pregnancy Using the Random Forest Approach - IEEE

[7] Prediction of Maternal Health Risk with Traditional Machine Learning Methods- (MTU Journal of Engineering and Natural Sciences

[8] A Semi-Supervised Machine Learning Approach in Predicting High-Risk Pregnancies in the Philippines- MDPI Diagnostics

[9] Early prediction of preeclampsia via machine learning- Original Research

[10] A Robust Machine Learning Predictive Model for Maternal Health Risk - IEEE