

MATERNAL HEALTH RISK PREDICTION BASED ON HEALTH CHECKUP USING MACHINE LEARNING APPROACHES

BY

**MD. MEHEDI HASAN
ID: 201-15-3516**

FINAL YEAR DESIGN PROJECT REPORT

This Report Presented in Partial Fulfillment of the Requirements for the
Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

MS. ISRAT JAHAN
Lecturer (Senior Scale)
Department of Computer Science and Engineering
Daffodil International University



DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

14 July 2024

APPROVAL

This Project titled “Maternal Health Risk Prediction Based on Health Checkup Using Machine Learning Approaches”, submitted by Md. Mehedi Hasan, ID No: 201-15-3516 to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 14 July, 2024.

BOARD OF EXAMINERS



Dr. Md. Ismail Jabiullah (MIJ)
Professor
Department of CSE
Daffodil International University

Board Chairman



Mr. Saiful Islam (SI)
Assistant Professor
Department of CSE
Daffodil International University

Internal Examiner 1



Mr. Afjal Hossan Sarower (AHS)
Sr. Lecturer
Department of CSE
Daffodil International University

Internal Examiner 2



Dr. Ahmed Wasif Reza (DWR)
Professor
Department of Computer Science and Engineering
East West University

External Examiner

DECLARATION

I hereby declare that, this project has been done by me under the supervision of Israt Jahan, Lecturer (Senior Scale), Department of CSE Daffodil International University. I also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Ms. Israt Jahan

Lecturer (Senior Scale)

Department of CSE

Daffodil International University

Submitted by:

Md. Mehedi Hasan

ID: 201-15-3516

Department of CSE

Daffodil International University

ACKNOWLEDGEMENT

First, I express my heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project/internship successfully.

I really grateful and wish my profound our indebtedness to Israt Jahan, Lecturer (Senior Scale), Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of “Research in Machine Learning” influenced me to carry out this research. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stage have made it possible to complete this project.

I would like to express my heartiest gratitude to Dr. Sheak Rashed Haider Noori, Professor & Head, Department of CSE, for his kind help to finish my project and also to other faculty member and the staff of CSE department of Daffodil International University.

I would like to thank my entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, I must acknowledge with due respect the constant support and patients of my parents.

ABSTRACT

Maternal health difficulties are currently one of the most difficult challenges in the world. Every year, many women die during pregnancy and after childbirth, which is a primary source of infant mortality. Maternal risk factors such as the mother's chronic illness, blood pressure, mental health, diet, and other medical care during pregnancy all play important roles. Pregnant women in remote locations confront several obstacles and challenges, including a scarcity of doctors, insufficient expertise, a lack of accessible clinics, infrastructural constraints, and transportation issues. The infant's poor health is mostly due to the mother's pregnancy, rather than any additional issues that may have occurred following childbirth. Using machine learning approaches, the study has predicted the maternal health risk level in previous due to avoid uncertain birth death or any inconvenience of a new born child. A variety of pre-trained advanced machine learning techniques were utilized in the study to find out the sustainable result. ANN, Ridge Classifier, SGD, XGBoost, Cat Boost, Random Forest, XGB, Decision Tree, and more algorithms were implemented. The recommended model was created, trained, and tested on the preprocessed dataset with the help of Hyper Parameter Tuning. The Cat Boost Classifier was the most accurate machine learning system for the study with a score of 97.4%.

Keywords— Maternal Health Risk, Machine Learning Approaches, Hyper Parameter Tuning, ANN, Ridge Classifier, AUC Score.

TABLE OF CONTENTS

Cotents	Page No
Board of Examiners	i
Declaration	ii
Acknowledgments	iii
Abstract	iv
List of Figures	vii
List of Tables	viii
CHAPTER 1: INTRODUCTION	1-4
1.1 Overview	1
1.2 Background and Present State	2
1.3 Problem Statement	2
1.4 Objectives	3
1.5 Scope and Limitations	3
1.6 Report Organization	3
1.7 Summary	4
CHAPTER 2: LITERATURE REVIEW	5-10
2.1 Overview	5
2.2 Related Works	5
2.3 Comparison between existing works	7
2.4 Open Issues	10
2.5 Summary	10
CHAPTER 3: METHODOLOGY/ REQUIREMENT ANALYSIS & DESIGN SPECIFICATION	11-15
3.1 Overview	11
3.2 Proposed Methodology/ System Design	12
3.3 Hardware/ Software Requirement	13
3.4 Project Management and Financial Analysis	14
3.5 Summary	15

CHAPTER 4: IMPLEMENTATION	16-19
4.1 Overview	16
4.2 Train Model/ Prototype Design	16
4.3 System Testing/ Model Evaluation	18
4.4 Summary	19
CHAPTER 5: RESULT AND ANALYSIS	20-25
5.1 Overview	20
5.2 Experimental/ Simulation Result	21
5.3 Performance/ Comparative Analysis	24
5.4 Summary	25
CHAPTER 6: IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY	26-28
6.1 Impact on Life	26
6.2 Impact on Society & Environment	26
6.3 Ethical Aspects	27
6.4 Sustainability Plan	27
6.5 Summary	28
CHAPTER 7: CONCLUSION AND FUTURE WORK	29-30
7.1 Conclusions	29
7.2 Further Suggested Works	29
7.3 Limitations/ Conflict of Interests	30
REFERENCES	31-32
APPENDIX	

LIST OF FIGURES

Figures	Page no
Figure 3.2.1: Workflow of Methodology.	12
Figure 3.2.2: Datasets	13
Figure 5.2.1: Roc Curve of Catboost	22
Figure 5.2.3: Confusion matrix of Catboost	22
Figure 5.2.4: Correlation Heat map.	23
Figure 5.2.6: Count plot for Outcome.	23
Figure 5.3.1: Accuracy Bar Chart.	24

LIST OF TABLES

Tables	Page no
Table 2.3.1: Comparative Analysis Table	7
TABLE 5.2.1: Accuracy for Train and Model.	21

CHAPTER 1

INTRODUCTION

1.1 Overview

The term "fatal maternal risk of complications" refers to a variety of issues that can have a major impact on a woman's and her children's health and occur during pregnancy, delivery, or the postnatal period. If some of these symptoms arise, the patient should seek medical attention right once to avoid death. Every year, approximately 585,000 women die around the world during childbirth, the postpartum period, or delivery, and nearly 50 million maternal and infant difficulties are reported, with approximately 300 million women affected by long-term and short-term medical issues associated with being pregnant, giving birth, and after delivery. The primary goal of this research is to predict health risk for maternal mortality in Bangladesh using machine learning techniques.

1.2 Background and Present State

Pregnancy health is a key measure of a country's general well-being and reflects the success of its healthcare services and regulations. Despite substantial improvement in other areas, maternal health remains a pressing concern in Bangladesh. In 2017, the World Health Organization (WHO) estimated that there were 196 maternal deaths per 100,000 live births in Bangladesh [1]. Although this figure has reduced from prior years, it underlines the necessity of ongoing efforts to guarantee healthy pregnancies and deliveries. Several factors impact the condition of maternal health in Bangladesh. Regional differences in maternal health outcomes continue, despite government attempts to improve healthcare infrastructure and services. A survey carried out by the Bangladesh Bureau of Statistics (BBS) in partnership with foreign organizations finds a large gap in maternal death rates between urban and rural areas [2].

Maternal health is a critical aspect of a woman's well-being during pregnancy, childbirth, and postpartum periods, and its impact extends beyond the individual to the broader community and society. The significant financial and human costs associated with maternal mortality underscore the importance of developing effective strategies to reduce these rates. Machine learning methods have shown promise in predicting maternal health risks by analyzing medical parameters such as maternal age, heart rate,

BMI, blood pressure, body temperature, and other relevant factors. By identifying early warning signs and categorizing risk factors, healthcare providers can take proactive measures to prevent complications and improve maternal health outcomes. The potential of machine learning to accurately predict risks, the significant impact of maternal deaths, and the need for targeted interventions drive the motivation behind maternal health risk prediction research.

1.3 Problem Statement

The research is driven by the urgent need to address the high rates of maternal mortality globally, particularly in developing countries. By utilizing machine learning methods to predict maternal health risks early, medical experts can intervene promptly to reduce complications and prevent maternal deaths. Factors such as age, blood disorders, blood pressure, blood glucose levels, body temperature, and heart rate directly impact pregnancy risks, highlighting the importance of early risk classification. Identifying and categorizing key risk factors through machine learning techniques can inform targeted interventions and evidence-based strategies to improve maternal health outcomes. The potential of machine learning to accurately predict risks, the significant impact of maternal deaths, and the need for tailored interventions are the primary motivations behind maternal health risk prediction research. This research aims to provide a proactive approach to maternal healthcare, enabling the early detection and management of pregnancy-related risks to ultimately reduce maternal mortality rates.

1.4 Objectives

- Identify the most accurate prediction of maternal health risk.
- Create public awareness about the maternal health and pregnancy death.
- Perceiving the maternal health risk prognosis results.
- Reducing the time and expense of discovering the maternal health risk.

1.5 Scope and Limitations

This study was part of a research project on the development and validation of machine learning models to predict risks for maternal health using complete medical checkup data. This covers anything from electronic health records to demographic material, responses, biometric statistics etc. In doing so, it will investigate different machine

learning algorithms to identify the best approach for risk prediction. The long-term goal is to develop a prognostic tool that will help health care professionals anticipate symptoms of pathologies early on based upon this knowledge, and thereby deliver personalized and timely interventions with the potential for better maternal/neonatal outcomes.

The study may face challenges related to data quality and completeness, as inconsistencies in health records can affect model accuracy. Additionally, there are concerns about model interpretability and the need for healthcare professionals to understand and trust the predictions. Ethical and privacy considerations related to the use of sensitive health data must be carefully managed. Furthermore, the generalizability of the models across different populations and healthcare settings may be limited, necessitating extensive validation efforts.

1.6 Report Organization

In the first chapter, I covered the introduction, motivation, study rationale, research questions, and expected outcomes. In a nutshell, chapter 1 elaborates on the project's inception. Chapter 2 contains information on Related Works, Comparative Analysis and Summary, Challenges, and other topics. In chapter 3, I discussed Research Methodology, Data Collection Procedure, and Statistical Analysis. Chapter 4 included experimental setup, results, and analysis. In Chapter 5, I discussed the impact on society, the impact on the environment, the Sustainability Plan, and so on. In Chapter 6, I reviewed the study and discussed the Future Scope and References.

1.7 Summary

For this work, a few numbers of essential machine learning algorithms are implemented. The best fit algorithm with highest accuracy can be used to predict the maternal health risk of each individual woman. Using this model, medicines might be deliberately assigned to high- risk women, possibly improving maternal health outcomes in Bangladesh. Machine learning algorithms are rapidly being researched for their ability to forecast health hazards for pregnant women, since these models have shown promise in healthcare contexts [23]. That is why novel ways to early identification and analysis of maternal health occurrences are necessary. This research suggests utilizing machine learning approaches to find out the best risk classifier for maternal mortality. As a consequence, I can figure out the cases by early detection and

the medical specialists will be able to make judgments that will allow for prompt patient treatment. This may help lower the likelihood of maternal and baby death during this time, as well as the social and economic impacts.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

Technology has made our lives so quick and simple, and we want everything at our fingertips. In this chapter, I will address relevant research or projects, as well as comparative analysis. I employed machine learning advanced classifiers and ensemble learning methodologies in the study to extract more accurate data. Machine learning or data mining techniques are more significant for making predictions in general.

Indeed, I can say that determining a woman's prenatal health risk status from medical exam records is my primary obligation. As if the patient could obtain exact information. Finally, knowing whether they are at high or low risk from the prediction system allows the patient to receive more effective treatment at the hospital.

2.2 Related Works

Raza et al. [4] developed an artificial neural network-based approach to improve maternal health prediction. The goal is to use health information to evaluate maternal health risks, and DT-BiLTCN is the recommended design. The support vector machine achieved an impressive 98% accuracy rate in this trial.

MUTLU et al. [5] proposed identifying and categorizing early risk factors for pregnancy disorders based on symptoms in another research. The authors used six different machine learning models to achieve their aim. The Decision Tree approach achieved the highest accuracy rate (89.16%) among all models examined. The paper "Ensemble Learning- Based Feature Engineering to Analyze Maternal Health and Predict Health Risks"

Ahmed et al. [6] investigated "IoT-Based Risk Level Prediction Model for Maternal Health Care in the Context of Bangladesh" to create a prediction model for maternal health utilizing data acquired from a variety of sources, including IoT devices, a web site, and Bangladeshi hospitals. The modified Decision Tree Algorithm was the most successful of the algorithms examined, with the best risk-level prediction accuracy. The claimed numerical figure for this precision is an incredible 97%.

Krupa et al. [7] employed Support Vector Machine (SVM) and empirical mode segmentation in their study of neonatal cardiotocography (CTG). The research

investigated ninety-nine randomly selected records, and the results indicated an 86% accuracy rate.

Nanda et al. [8] investigated the role of biomarkers and specific maternal characteristics in predicting the development of diabetes in the first trimester of pregnancy. Using the found variables, the researchers built a prediction model using logistic regression. The findings revealed that the model predicted properly 74.1 percent of the time.

Pawar et al. [9] used machine learning approaches to detect maternal health risk factors. The researchers say that these computational methods may be used to identify and assess maternal health hazards. Surprisingly, they stated that their model beat conventional machine learning techniques. According to their findings, the constructed model has a 70.21% accuracy rate.

Chaudhuri and Mandal [10] used NFHS-3 data to explore the complicated patterns of maternal and child mortality in West Bengal. They conducted extensive analysis to identify socioeconomic characteristics that influence the use of maternal health inputs. Using a mixed-methods approach, the study considered both endogenous and exogenous estimates, providing comprehensive insights into the determinants of prenatal care, birthplace choice, and their subsequent influence on infant mortality.

Assaduzzaman et al. [11] conducted a groundbreaking study on the timely prediction of maternal health risk factors. To develop effective prediction models, the researchers used a preprocessed dataset and a variety of machine-learning methods. Random Forest was the most effective algorithm among the ones examined in terms of performance.

Elia and Ayungo [12] investigated the key characteristics influencing pregnant women's acquisition of maternal health information. Their cross-sectional study, which included surveys and focus group talks with 132 pregnant women and 8 healthcare professionals, discovered a pattern in participants' total knowledge of maternal health data.

Kenny et al. [13] sought to provide a prediction for pregnant mothers with preeclampsia. They used genetic programming on 97 plasma samples to validate the chemical patterns that distinguish the fluid of pre-eclampsia patients.

Yarlapati et al. [14] used machine learning to assess the likelihood of preterm delivery during a pregnancy, which is a crucial metric of the fetus's health. They were quite confident in declaring with 96.77% accuracy that the Bayes algorithm with the lowest fail rate identifies the babies' condition based on clinical information.

The authors of [15] presented a model to predict hypertension during the first trimester of pregnancy. The proposed model achieved an impressive 90% accuracy, according to the provided data. Furthermore, the model had a low 5% false positive rate. For forecasting blood sugar levels, matching indicators, and high blood pressure.

Rai et al. [16] evaluated the parameters linked with an increased risk of maternal and newborn mortality as a result of pregnancy-related complications. The researchers utilized a questionnaire to get feedback from subject matter experts for their three assessments. Fourteen criteria were selected for the study from a dataset including information from 117 expectant moms. To estimate risk levels, the researchers used feed-forward artificial neural network (ANN) and Naive Bayes (NB) techniques. The findings revealed that the ANN algorithm had an accuracy rate of 80%, whereas the Naive Bayes technique had a slightly lower accuracy rate of 70%.

The authors [17] used a hybrid technique that included logistic regression, K-Nearest Neighbors, dimension reduction with multiple variables, and support vector machines to predict blood sugar levels, matching symptoms, and high blood pressure.

The obstacles of maintaining maternal health persist in Bangladesh's distinct socioeconomic and healthcare environments. Understanding these specific challenges necessitates a review of the existing maternal health literature. Although scholars have sought to emphasize the effective implementation of healthcare in a variety of foreign settings [18], there is a notable lack of research focusing on Bangladesh's specific standpoint.

2.3 Comparison Between Existing Works

Most prior studies on the prediction of maternal health risk attempted to experiment with a few pre-trained data mining algorithms. Many of them employed the minimal amount of available data and chose factors that were occasionally unrelated to the diagnosis of maternal health risk. A handful of them also worked with survey results and open-source datasets. Some writers are also skeptical about the preprocessing methods.

TABLE 2.3.1: Comparative Analysis Table

Authors	Goal	Method/Model(s) Used	Best Performing Model/Algorithm	Accuracy Rate
Raza et al. [4]	Improve maternal health prediction using health information	Artificial Neural Network (ANN), DT-BiLTCN, Support Vector Machine (SVM)	Support Vector Machine (SVM)	98%
MUTLU et al. [5]	Identify and categorize early risk factors for pregnancy disorders	Six different machine learning models	Decision Tree	89.16%
Ahmed et al. [6]	Predict maternal health risks using IoT data in Bangladesh	Various machine learning models, including modified Decision Tree	Modified Decision Tree	97%
Krupa et al. [7]	Analyze neonatal cardiotocography (CTG)	Support Vector Machine (SVM) and empirical mode segmentation	Support Vector Machine (SVM)	86%
Nanda et al. [8]	Predict development of diabetes in the first trimester of pregnancy	Logistic regression	Logistic Regression	74.1%
Pawar et al. [9]	Detect maternal health risk factors	Various machine learning models	Custom Model	70.21%
Chaudhuri and Mandal [10]	Analyze maternal and child mortality in West Bengal using socioeconomic data	Mixed-methods approach	Not specified	Not specified

Assaduzzaman et al. [11]	Predict maternal health risk factors	Various machine learning models	Random Forest	Not specified
Elia and Ayungo [12]	Investigate key characteristics influencing maternal health information acquisition	Cross-sectional study, surveys, and focus groups	Not specified	Not specified
Kenny et al. [13]	Predict preeclampsia using genetic programming	Genetic programming	Not specified	Not specified
Yarlapati et al. [14]	Assess likelihood of preterm delivery using machine learning	Bayes algorithm	Bayes algorithm	96.77%
Authors [15]	Predict hypertension during the first trimester of pregnancy	Predict hypertension during the first trimester of pregnancy	Not specified	90%
Rai et al. [16]	Evaluate risk parameters linked to maternal and newborn mortality	Feed-forward Artificial Neural Network (ANN), Naive Bayes (NB)	ANN	80% (ANN), 70% (Naive Bayes)
Authors [17]	Predict blood sugar levels, symptoms, and high blood pressure	Hybrid method (Logistic Regression, KNN, SVM, etc.)	Not specified	Not specified

Using Google Colab, I will combine common and popular classifiers such as Support Vector Classifier (SVC), Decision Tree, and Naive Bayes with advanced machine

learning algorithms such as Stochastic Gradient Descent, Ridge Classifier, Extreme Gradient Boosting (XGBoost), and CatBoost to predict maternal health risk and obtain an accurate classifier for this study. I also used the Artificial Neural Network (ANN) in the experiment, and it worked great. More crucially, I have 1524 occurrences of local data in my dataset. The experiment collects the required variables for identifying maternal health risk. Finally, I observed that the CatBoost algorithm had the highest accuracy for forecasting maternal health risk based on various medical exam data.

2.4 Open Issues

In the medical profession, there is little opportunity to collect data. Due to privacy concerns, the majority of institutes do not share their patients' medical checkup records or data. I visited several mother and child centers, hospitals, and clinics to obtain medical checkup data, but many of them refused to offer the information. Following that, I found a private clinic and diagnostic center that agreed to supply data for the purpose of my research.

On the other hand, to develop a model using machine learning techniques, various obstacles may occur, such as Collect real-time medical data and Understanding machine learning techniques.

2.5 Summary

Conclusion In essence, the adequacy of accuracy rates evidenced in this literature review shows great promise for the ability to predict maternal health risk with machine learning models using diverse data sources. Nevertheless, numerous challenges arise: data quality and reliability, model interpretability; ethical issues; integration into clinical care. More broadly applicable models that include different populations and data in real-time are also required. Future research is needed to tease these open issues in order for machine learning tools that can be implemented internationally to see extensive use and save more mothers' lives.

CHAPTER 3

METHODOLOGY/ REQUIREMENT ANALYSIS & DESIGN

SPECIFICATION

3.1 Overview

The study "Maternal Health Risk Prediction Based on Health Checkup Using Machine Learning Approaches" provides an overview of the research on maternal health risk prediction. Viral infections impact millions of individuals globally, and their symptoms are frequently misdiagnosed until they become serious. Effective management and the avoidance of problems depend heavily on early detection and intervention. The clinical examination and laboratory tests that are the mainstay of traditional diagnostic approaches are not always reliable in predicting maternal health concerns or determining the severity of those risks. With the use of machine learning (ML) and artificial intelligence (AI) advancements, this research attempts to create predictive models that will help with maternal health early detection and severity evaluation.

In a nutshell, this strategy uses machine learning techniques to create prediction models that identify people who may already be at high risk for maternal health issues or who may already be at risk but have not yet received a diagnosis. In order to forecast the possibility or severity of maternal health issues, the study examines the viability, accuracy, and efficacy of several machine learning algorithms that can evaluate pertinent data such as medical history, environmental factors, symptoms, and potentially even genetic predisposition. The comparative part assesses how well various algorithms perform.

Informed judgments made by public health officials and healthcare experts could lead to better patient outcomes, lower healthcare costs, and early interventions. This research is utilizing machine learning approaches to construct models of maternal health risk. There is a lot of potential to improve patient care and healthcare delivery by incorporating machine learning algorithms into public health initiatives and maternal health diagnoses. This study aims to solve current issues and open the door for more precise, effective, and individualized management of maternal health by utilizing the potential of data-driven approaches.

3.2 Proposed Methodology/ System Diagram

There are numerous important steps in the suggested machine learning-based maternal health risk prediction method. To ensure data quality and applicability for training models, secondary datasets composed of patient data and risk level are first gathered and pre-processed. It manages missing values, categorical variable encoding, and scaling of numerical features. Subsequently, feature selection approaches are employed to ascertain the relative importance of key maternal health predictors. The pre-processed dataset is used to train several machine learning algorithms after feature selection. To optimize and prevent model performance issues, cross-validation and hyper parameter adjustment are employed.

Over-adaptation First, preprocessed datasets comprise patient medical records with variables like age, blood pressure, systolic blood pressure, medical history, and blood test results are gathered. This include encoding categorical variables, normalizing features, and cleaning the data. To find the most pertinent predictors, feature selection is then carried out utilizing statistical approaches and dimension reduction methods. K-fold cross-validation and hyper parameter tuning are used in the training and validation of several machine learning methods. A confusion matrix is used to illustrate performance, and metrics including accuracy, precision, recall, F1-score, and AUC-ROC are used to evaluate the models. Attribute importance scores are examined in order to comprehend the model. The flowchart is given below:

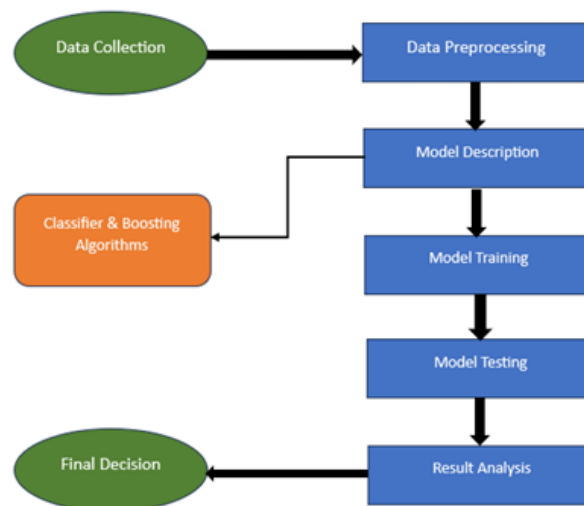


Figure 3.2.1: Workflow of Methodology.

Datasets

Patient data gathered from the closest hospital is included in the dataset used in this investigation. A wide range of clinical, diagnostic, and demographic data pertinent to the prediction of maternal health are included in the dataset. Important factors gathered for every patient's record.

Age	Systolic	Diastolic	BS	BT	BMI	PC	PD	GD	MH	Heart	Risk
23	100	60	9	100	18	1	1	0	1	80	High
24	110	60	6.5	98	21.2	0	0	0	0	70	Low
44	140	100	14	98	25.3	0	1	0	1	88	High
37	110	70	7.5	98	23	0	0	0	0	68	Low
34	100	60	14	98	30	1	0	1	1	74	Low

Figure 3.2.2: Datasets

The dataset includes 12 attributes, which correlate to different patient features, and 1524 patient records, each of which is displayed as a row. Data collecting methods guarantee the confidentiality and anonymity of patient information by abiding by ethical standards and patient privacy laws. To ensure data quality and consistency for machine learning analysis, preprocessing activities including handling missing values, encoding categorical variables, and scaling numerical features were carried

3.3 Hardware/ Software Requirement

For Sturdy hardware and software are necessary for the implementation of a machine learning system for dengue prediction. To do complex computations, a high-performance computer with a multi-core CPU, SSD storage, at least 16 GB of RAM, and a dedicated GPU is needed. Scalability may require additional cloud resources, such AWS or Google Cloud. Programming languages like Python and a compatible operating system (Windows, MacOS, or Linux) are essential in terms of software. Pandas for data manipulation, Skit-Learn for machine learning algorithms, and TensorFlow or PyTorch for model development are important libraries and frameworks. To guarantee efficient development, collaboration, and deployment processes, tools for version control (Git), data management (SQL/NoSQL databases), and deployment (Docker, Flask) are also crucial.

Hardware

- High-end computers.
- Server infrastructure.
- Advanced devices.

Software

- Operating system.
- Anaconda / Jupyter Notebooks.
- Google Colabs.
- Data management and processing tools.
- Deployment and monitoring.
- Security and compliance tools.

3.4 Project Management and Financial Analysis

To facilitate the study's smooth progression and implementation, the system requirements are stated below.

- A laptop/desktop with at least 4 GB RAM.
- Anaconda 3 software.
- Jupyter Notebook IDE/ Google Colab.

3.5 Summary

Preprocessing, model selection, evaluation, and data collecting are all part of the methodical methodology used in this study, "Maternal Health Risk Prediction Based on Health Checkup Using Machine Learning Approaches." To ensure a thorough dataset, data will be gathered via clinician reports, patient surveys, and electronic health records. In order to improve model accuracy, preprocessing procedures include resolving missing values, standardizing data, and choosing pertinent features. For the classification challenge, eight machine learning algorithms have been chosen based on their different strengths.

To guarantee robustness and avoid over fitting, a cross-validation technique will be used for both training and validation of each model. A number of metrics, including F1-score, recall, accuracy, and precision, will be used to assess performance. Data security and patient privacy are two ethical issues that are central to the design. The ultimate goal of this all-encompassing strategy is to create a therapeutically valuable,

interpretable, and dependable model for predicting maternal health, which will enhance patient outcomes and care efficiency.

CHAPTER 4

IMPLEMENTATION

4.1 Overview

The steps done in practice to create machine learning models for dengue forecasting are covered in the implementation section. We will now reveal some machine learning algorithms that were considered for this assignment. We'll also concentrate on the pre-processing methods required to make sure the data is organized, clean, and appropriate for the algorithm of choice. This section will go into more depth about the training procedure used to create the model, namely how the model describes dengue-related patterns. Lastly, We'll talk about a few assessment criteria that we used to gauge how well the model predicted the occurrence of dengue. A road map for comprehending the technological intricacies needed in developing machine learning models for dengue forecasting is given in this overview.

4.2 Train Model/ Prototype Design

Predictive models are developed and validated through a number of crucial phases in the use of "Maternal Health Risk Prediction Based on Health Checkup Using Machine Learning Approaches". First, a comprehensive preprocessing is performed on the dataset, which consists of 1524 patient records with 12 characteristics. These include selecting features to improve model performance, normalization, and handling missing data. Based on initial evaluation measures and their fit for the classification problem, eight machine learning methods were chosen.

Train Model:

Based on their suitability for the classification task and initial evaluation metrics, eight algorithms are selected: Artificial Neural Networks (ANN), Naive Bayes, Decision Trees, Support Vector Classifier (SVC), Stochastic Gradient Descent (SGD), Extreme Gradient Boosting, CatBoost, and the Ridge Classifier. To preserve the distribution of classes, the dataset is split into training and test sets, usually using stratified sampling approaches. Each algorithm is trained on various subsets of the training data using cross-validation techniques like cross-validation, and their performance is assessed over a number of rounds.

Through training, the algorithms develop the ability to use the training data to generate predictive models that can differentiate between individuals who pose a high risk and those who do not. Standard metrics like accuracy are used to evaluate the performance of the models in order to determine how well each algorithm generalizes to new, unseen data.

To further maximize performance, each algorithm's unique hyper parameters are fine-tuned during the training process. By avoiding over fitting, this iterative procedure guarantees that the models reach high accuracy and robustness. The outcomes of this training phase set the stage for the validation and subsequent application of machine learning models in clinical practice, with the aim of enhancing patient care outcomes associated with maternal health risks and raising the diagnostic accuracy.

Prototype Design:

In order to incorporate the trained models into clinical practice, the prototype design for "Maternal Health Risk Prediction Based on Health Checkup Using Machine Learning Approaches" focuses on creating an intuitive user interface. In addition to ensuring accessibility for medical practitioners, the design seeks to optimize the patient data-driven prediction process.

User Interface: A graphical user interface (GUI) on the prototype makes it simple for medical professionals to enter patient data. Clear prompts for entering pertinent medical data, such as symptoms, medical history, and demographic data, are incorporated into the interface's straightforward design.

Data Integration: The prototype easily interfaces with standalone data input techniques or current electronic health record (EHR) systems. This guarantees that medical professionals can safely and effectively obtain patient data for forecasting needs.

Prediction Engine: Extreme Gradient Boosting (XGBoost), CatBoost, the Ridge classifier, and Decision Tree are the machine learning models that have been trained behind the interface of the prototype. These algorithms can forecast a patient's chance of experiencing maternal hazards by analyzing the data that is input.

Output and Interpretability: The prototype indicates the expected probability or classification of maternal health risk in a way that is easy for physicians to understand.

Anticipations can be accompanied by visual aids for understanding and decision-making, such as charts or risk level.

Ethical considerations: Ethical concerns around patient data protection and model openness are critical throughout the entire design process. The prototype has features to help clinicians understand model predictions while also guaranteeing compliance with data protection rules.

Testing and Validation: The prototype is put through extensive testing before being put into use to make sure it is accurate, dependable, and useful in actual clinical settings. Predictive models' effectiveness and generalizability are further validated via validation against separate datasets or via clinical trials.

Iterative Improvement: Iterative upgrades derived from enhancements made to performance metric prototypes and ongoing input from healthcare professionals. Over time, this iterative method seeks to increase user satisfaction, prediction accuracy, and overall clinical utility.

By concentrating on these design concepts, the prototype seeks to improve patient care outcomes in maternal health risk projections and close the gap between machine learning research and real-world healthcare applications.

4.3 System Testing/ Model Evaluation

To assure accuracy, dependability, and usability in clinical practice, "Maternal Health Risk Prediction Based on Health Checkup Using Machine Learning Approaches" is being implemented with thorough system testing and model validation.

Testing Framework: A thorough testing protocol is developed to assess the performance of the prototype. Unit testing of distinct parts, such as the user interface, machine learning algorithms, and data pretreatment modules, is included in this. The smooth operation of various system modules, such as the data input, prediction engine, and output visualization, is ensured via integration testing.

Performance Metrics: Common performance measures, including accuracy, are used to evaluate the efficacy of the prediction models. Based on the input data, these measures offer quantitative insights into how successfully the models classify patients into high-risk and low-risk groups.

Cross-Validation: Cross-validation is one of the cross-validation approaches used to test the robustness and generalizability of the models. Training and testing are carried

out iteratively to reduce overfitting and make sure the models function effectively on data that hasn't been seen by dividing the dataset into several subsets.

Validation Against Baselines: The efficacy of the machine learning models is evaluated in relation to baseline techniques or accepted clinical standards. The models' superiority in terms of accuracy and predictive capacity is supported by this comparative study.

Real-World Simulation: The performance of the prototype under various settings can be further assessed by using datasets and simulated real-world scenarios. During this testing step, the predictive models' adaptability to various patient demographics, symptom profiles, and clinical contexts is evaluated.

Clinical Validation: Clinical validation of the prototype is facilitated by collaboration with healthcare experts. Physicians assess the predictions made by the prototypes against the actual outcomes of their patients, offering qualitative comments on the prototypes' usability, interpretability, and therapeutic significance.

Ethical Considerations: Ethical concerns of patient privacy, data security, and model openness are critical throughout the testing and evaluation phase. When it comes to the use of healthcare data, steps are taken to guarantee compliance with legal requirements and ethical norms.

This implementation phase seeks to evaluate the efficacy and usability of machine learning-based maternal health risk prediction models in real-world clinical settings through extensive system testing and model evaluation. The evaluations yield valuable insights that aid in the improvement of patient care outcomes through model refinement and accuracy.

4.4 Summary

The steps involved in creating a machine learning model to forecast the hazards to a mother's health are described in this section. We go into detail about the machine learning algorithm that was chosen, the methods for cleaning and preparing the data before analysis, and the model-building training procedure. Furthermore, we delineate the assessment measures employed to appraise the model's efficacy in risk prediction. A general review of the technical issues of creating machine learning models for forecasting maternal health is given in this section.

CHAPTER 5

RESULT AND ANALYSIS

5.1 Overview

Insightful findings about the performance of eight algorithms were obtained from this study on "Maternal Health Risk Prediction Based on Health Checkup Using Machine Learning Approaches": Artificial Neural Networks (ANN), Naive Bayes, Decision Trees, Support Vector Classifier (SVC), Stochastic Gradient Descent (SGD), Extreme Gradient Boosting, CatBoost, and the Ridge Classifier. Next, the dataset is divided into testing and training sets. Accuracy and area under the receiver operating characteristic curve are two examples of suitable evaluation measures that are used to assess the performance of the model (AUC-ROC). To evaluate the effectiveness of several algorithms and choose the best technique for predicting the risk to a mother's health, comparative analysis is conducted. To guarantee the accuracy and applicability of the findings, validation methods such as external validation and sensitivity analysis are used. The study intends to create accurate and trustworthy predictive models for maternal health forecasts based on this experimental design, which would enhance patient outcomes and healthcare decision-making. According to the study Catboost outperformed the other algorithms in terms of accuracy, achieving a remarkable 97.4%. The resilience and appropriateness of Catboost for medical prediction tasks are demonstrated by their abilities to classify patients into normal and afflicted groups. XGBoost also fared well, obtaining a 96.9% accuracy rate, indicating its capacity to manage intricate variables and data structures.

These findings imply that machine learning algorithms might greatly enhance the precision and effectiveness of high- and low-risk prediction, giving medical professionals useful instruments to enhance patient care plans and diagnostic procedures. Additional examination reveals the necessity of ongoing model validation and improvement for various patient populations.

5.2 Experimental/ Simulation Result

In this section, eight algorithms are compared. The performance of classification comes first. The overall metrics of those models are then examined in order to gather, characterize, pinpoint probable causes, and pinpoint areas where outcomes could be improved. First, for each machine learning method under consideration, the section

summarizes performance indicators like as accuracy and area under the receiver operating characteristic curve (AUC-ROC). The performance of various algorithms is then compared using comparative analysis in order to determine the best technique for predicting maternal health risk. To further demonstrate the effectiveness of the models, additional visualizations like ROC curves and confusion matrices can be offered. Discussed are the advantages and disadvantages of each method, emphasizing elements like computational efficiency, interpretability, and forecasting accuracy.

Sensitivity analysis is used to evaluate the predictability of the model under various scenarios, and if external validation methods are available, they can be used to evaluate the model's performance on independent datasets. All things considered, the analyses and outcomes of the experiments offer insightful information about how well machine learning models perform when predicting maternal health risk using secondary datasets, which will guide future research and clinical decision-making procedures.

TABLE 5.2.1: Accuracy for Train and Model.

Architecture	Model Accuracy
ANN	96.3%
Naïve Bayes	95%
Decision Tree	95.9%
Support Vector Classifier	96%
Stochastic Gradient Descent	95.9%
Extreme Gradient Boosting	96.9%
CatBoost	97.4%
Ridge Classifier	95.7%

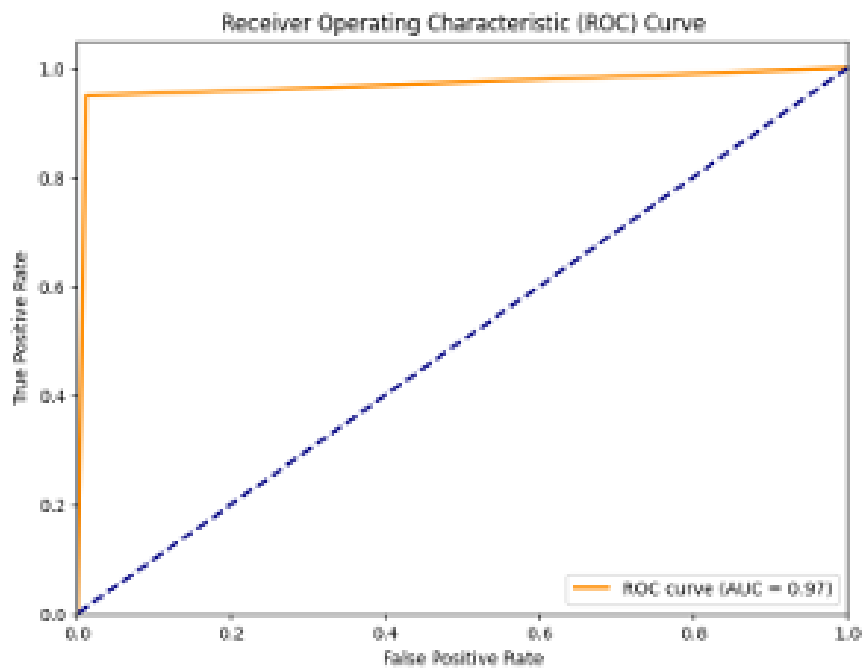


Figure 5.2.1: Roc Curve of Catboost

Confusion Matrix after Selecting Feature:

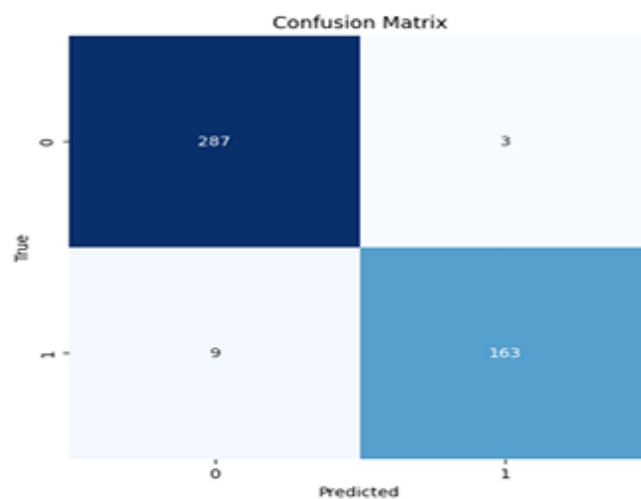


Figure 5.2.3: Confusion matrix of Catboost

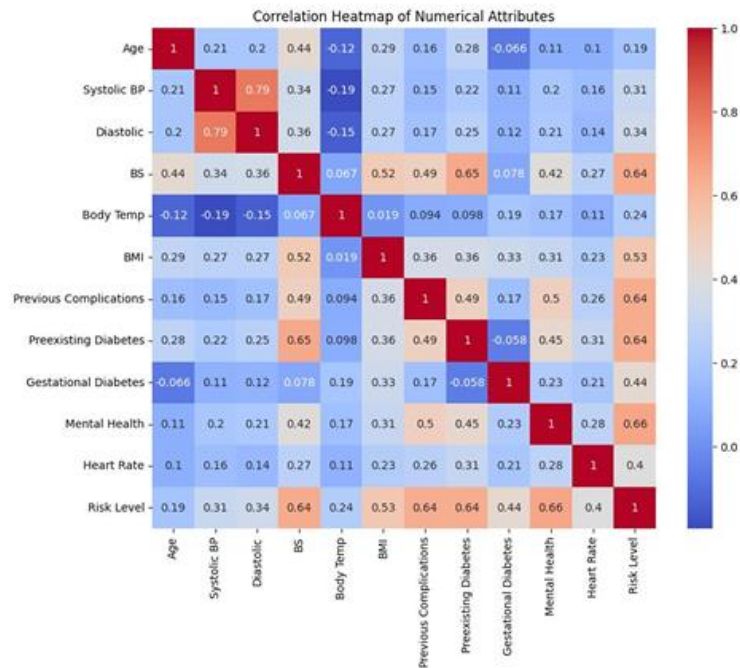


Figure 5.2.4: Correlation Heat map.

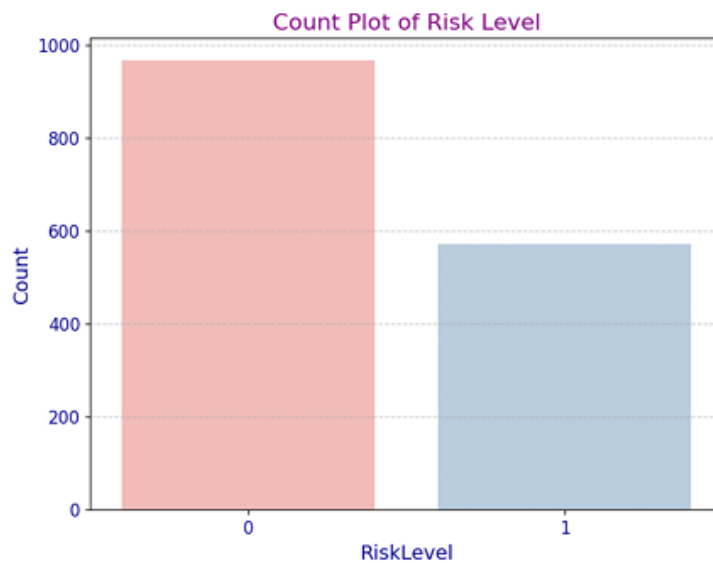


Figure 5.2.6: Count plot for Outcome.

The high ratio of true positives to true negatives suggests that the model performs well overall. Still, a tiny fraction of false positives and false negatives happen. While false positive results may result in patients receiving unneeded therapy who do not genuinely pose a significant risk, false negative results may make patients wait longer for critical treatment. Remember that this is just a small sample of data, and that the model's performance can change depending on the dataset that was used to train it. Moreover, the confusion matrix presents only a portion of the story. Extra measures that can provide further insight into the effectiveness of the model.

5.3 Performance/ Comparative Analysis

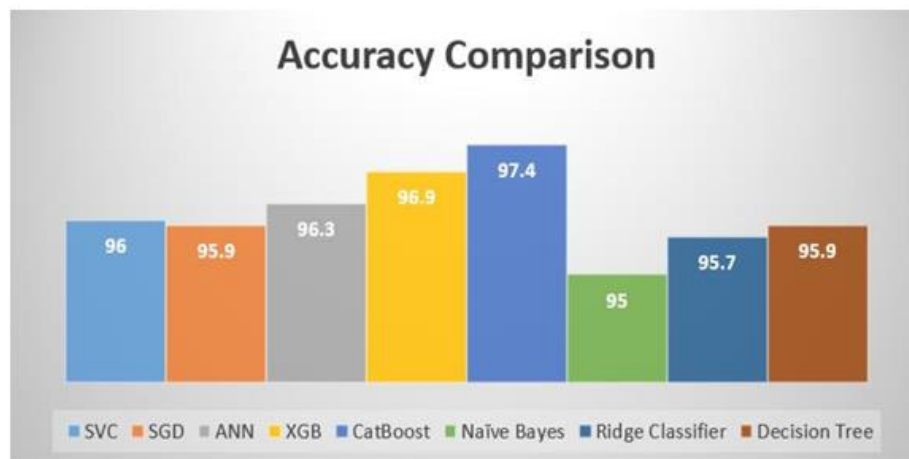


Figure 5.3.1: Accuracy Bar Chart.

Several machine learning algorithms, including Catboost with 97.4%, XGBoost with 96.4%, and ANN with 96.3%, achieve high accuracy in this section of the thesis paper "Maternal Health Risk Prediction Based on Health Checkup Using Machine Learning Approaches." Several points can be highlighted to contextualize and analyze these results. The conversation can first center on contrasting each algorithm's performance and highlighting its advantages and disadvantages for predicting maternal health risks. For instance, random forest's performance can be ascribed to its capacity to handle high-dimensional data and grasp intricate relationships between features, while catboost's high accuracy can be related to its ability to construct nonlinear decision boundaries effectively. Likewise, the conversation might delve into the balance between interpretability and accuracy, pointing out that even if XGBoost attains greater accuracy, its more straightforward decision-making procedures might provide greater interpretability. The potential use of machine learning algorithms to help healthcare providers advise targeted interventions or preventive measures is also discussed. In addition, possible sources of bias or confounding that could impair model performance and the generalizability of results to other patient groups or healthcare environments could be discussed. Overall, the talk offers important context for understanding the importance of the high accuracy in predicting maternal health that several machine learning algorithms achieve, as well as research implications.

5.4 Summary

The outcomes of the machine learning model created for forecasting maternal health are shown in the Results and Analysis section. Here, we go over the model's performance using a few chosen assessment metrics. In order to evaluate the model's predictive accuracy for maternal health risk, accuracy measures are presented. To learn more about the advantages and disadvantages of the model, we shall examine the data in more detail. This could involve looking into what influences the model's performance, figuring out where it could be improved, or using the model's predictions to visualize the decision-making process. The performance of machine learning models created for forecasting maternal health is examined in the Results and Analysis section. We examined eight different algorithms.

With an accuracy of 97.4%, Catboost was the most accurate of these algorithms, followed by XGBoost (96.9%) and ANN (96.3%). This suggests that the best model for correctly predicting the presence of risks in our dataset was catboost. Subsequent examination will go more deeply into each model's performance. We will investigate variables that might have an impact on the outcomes, like the intricacy of the data or certain traits of each program. Furthermore, we might think of visualization methods to comprehend the predictions made by each model and pinpoint possible areas for development. The purpose of this part is to give a thorough understanding of how well the models anticipate high-risk and low-risk events.

CHAPTER 6

IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

6.1 Impact on Life

This study may have a tremendous influence on life, influencing many aspects of public health, healthcare, and personal well-being. Early discovery of maternal health risks can lead to rapid medical intervention, potentially lowering the risk of pregnancy and newborn mortality. This research might lead to improved medical processes, regulations, and recommendations for identifying and treating maternal health issues. By utilizing machine learning algorithms to predict health risks during pregnancy, this research empowers healthcare providers to intervene early, leading to better health monitoring and treatment for expectant mothers. The implementation of predictive models enhances healthcare practices, ensuring timely interventions and personalized care for pregnant women, ultimately contributing to healthier pregnancies and safer childbirth experiences. Moreover, by addressing maternal health risks proactively, this research not only benefits individual women but also has broader societal implications, such as reducing the economic burden associated with maternal mortality, improving overall community health, and promoting maternal well-being as a fundamental pillar of society. On the other hand, the study can assist improve healthcare technology, especially if it promotes the use of new advancements such as machine learning.

6.2 Impact on Society and Environment

The impact of maternal health risk research on the environment is significant. By understanding the environmental factors affecting maternal health, such as pollution, chemicals, and climate change, researchers can highlight the interconnectedness between environmental health and maternal well-being. This research sheds light on how environmental exposures can impact pregnancy outcomes, maternal morbidity, and mortality, emphasizing the need for policies and interventions to address these risks. Furthermore, the identification of environmental stressors like air pollution, noise, and inadequate green spaces underscores the importance of creating healthier built 28 ©Daffodil International University environments to support maternal and infant health. Ultimately, maternal health risk research not only improves healthcare practices but also advocates for sustainable environmental policies that benefit both maternal health and the well-being of communities at large.

6.3 Ethical Aspects

The study ensuring the adequacy of informed consent, considering paternal involvement, balancing risks for both the mother and fetus, and promoting access to research and treatment. These ethical considerations are crucial in safeguarding the rights, autonomy, and well-being of pregnant women, especially those living with HIV, and their unborn children. Respecting the core ethical principles of Beneficence, Respect for Persons, and Justice is essential in conducting research with pregnant women to ensure their safety, understanding, and involvement in decision-making processes. Balancing the risks involved, addressing disclosure concerns, and promoting inclusivity in research participation are key ethical imperatives in maternal health risk prediction research. However, the technological involvement will occur.

6.4 Sustainability Plan

The predictive models are integrated into normal maternity care by forming collaborations with healthcare providers, legislators, and community organizations. Embracing digital platforms expands access to risk assessment tools and encourages continual data collecting for model modification. Furthermore, investing in capacity building and education allows healthcare practitioners to successfully use these tools. Long-term financing commitments and stakeholder participation are essential for fostering an innovative and adaptable culture in maternity healthcare practices. People must be aware of the significance of maternal health risk finding and usage of the system. Users can gain benefits from the system. Given these considerations, the system's sustainability strategy was fulfilled

6.5 Summary

The application of machine learning to the prediction of maternal health has profound effects on sustainability, the environment, and society. Socially, this technology helps distant populations and improves healthcare equity and accessibility. It lowers the frequency of high-risk cases, lessens the financial strain on patients and healthcare systems, and promotes a proactive attitude toward health. Reducing hospital stays and critical care helps the environment by using less resources and medical waste, which promotes sustainability. As a result of fewer hospitalizations, throwaway medical supplies are used less frequently, supporting efforts to preserve the environment. By decreasing the need for substantial physical infrastructure and travel, the adoption of

digital health solutions lowers the healthcare industry's carbon footprint. Overall, including machine learning into the process of predicting the danger to a mother's health promotes an effective and sustainable conclusion that benefits the environment and society.

CHAPTER 7

CONCLUSION AND FURTHER WORK

7.1 Conclusions

In today's world, we are all dependent on technology. We cannot survive a day without the use of technology. However, technology and appliances offer both advantages and downsides. In the world of technology, the implementation of machine learning is rather difficult. The medical and technological sectors are interdependent. Today, we observe a revolutionary technology participation in the medical business. The proposed method predicts maternal health risks using data from patient medical checkups. To predict maternal health risk, I evaluated the effectiveness of several classifiers and boosting strategies. When comparing the accuracy of different machine learning algorithms, CatBoost performs remarkably well across a variety of criteria, whereas Naive Bayes has the lowest accuracy. Python and Google Colab were used to conduct the research. Pregnancy death can result from a significant risk of maternal health. This automated technique might be a useful tool in raising awareness of whether maternal health risks are high or low. This tool can produce the most rapid and accurate prediction results in the shortest amount of time and at no cost.

7.2 Further Suggested Works

The world operated on the basis of artificial intelligence. Machine learning is part of artificial intelligence. Our daily lives provide several challenges. Machine learning may 30 ©Daffodil International University significantly improve problem-solving. Machine learning has a large variety of algorithms. The study may be expanded with new data, data mining tools, and improved machine learning algorithms. In the near future, I want to develop an Android app and web-based software that can quickly assess maternal health risks and recommend simple preventative steps to consumers.

7.3 Limitations/ Conflict of Interests

Despite the significant contributions of this research, several limitations must be acknowledged. The dataset used was relatively small, which may impact the generalizability of the findings. The real-life implementation was confined to a small-scale pilot project, limiting the broader applicability of the results. Specific challenges were encountered in addressing diverse maternal health issues due to limited access to

comprehensive data and the complexity of these concerns. Additionally, geographical constraints and resource limitations, including time, funding, and technical infrastructure, may have affected the depth and breadth of the research. Addressing these limitations in future studies will be crucial for enhancing the validity and applicability of the findings.

REFERENCES

- [1] “Global Health Observatory,” Who.int, 2019. <https://www.who.int/data/gho>. “Bangladesh Bureau of Statistics,” Dghs.gov.bd, 2023.
- [2] A. Rajkomar, J. Dean, and I. Kohane, “Machine Learning in Medicine,” *New England Journal of Medicine*, vol. 380, no. 14, pp. 1347-1358, Apr. 2019.
- [3] A. Raza, H. U. R. Siddiqui, K. Munir, M. Almutairi, F. Rustam, and I. Ashraf, “Ensemble learning-based feature engineering to analyze maternal health during pregnancy and health risk prediction,” *PLOS ONE*, vol. 17, no. 11, p. e0276525, Nov. 2022.
- [4] H. B. Mutlu, F. Durmaz, N. Yucel, Cengi Ie., and M. Yildirim, “Prediction of Maternal Health Risk with Traditional Machine Learning Methods,” *NATURENGS*, vol. 4, no. 1, pp. 16–23, Jun. 2023.
- [5] N. Krupa, M. MA, E. Zahedi, S. Ahmed, and F. M. Hassan, “Antepartum fetal rate feature extraction and classification using empirical mode decomposition and support vector machine,” *Biomedical Engineering OnLine*, vol. 10, no. 1, p. 6, 2011.
- [6] S. Nanda, M. Savvidou, A. Syngelaki, R. Akolekar, and K. H. Nicolaides, “Prediction of gestational diabetes mellitus by maternal factors and biomarkers at 11 to 13 weeks,” *Prenatal Diagnosis*, vol. 31, no. 2, pp. 135–141, Dec. 2010.
- [7] L. Pawar, J. Malhotra, A. Sharma, D. Arora, and D. Vaidya, “A Robust Machine Learning Predictive Model for Maternal Health Risk,” *2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC)*, Aug. 2022.
- [8] S. Chaudhuri and B. Mandal, “Predictive behaviour of maternal health inputs and child mortality in West Bengal – An analysis based on NFHS3,” *Heliyon*, vol. 6, no. 5, pp. e03941–e03941, May 2020.
- [9] M. Assaduzzaman, A. A. Mamun, and M. Z. Hasan, “Early Prediction of Maternal Health Risk Factors Using Machine Learning Techniques,” *IEEE Xplore*, Jan. 01, 2023.
- [10] E. F. Elia and J. Ayungo, “Socio-demographic influence on the pregnant women’s comprehension of maternal health information in Tanzania,” *Heliyon*, vol. 9, no. 12, p. e22448, Dec. 2023.
- [11] L. C. Kenny, W. B. Dunn, D. I. Ellis, J. Myers, P. N. Baker, and D. B. Kell, “Novel biomarkers for pre-eclampsia detected using metabolomics and machine learning,” *Metabolomics*, vol. 1, no. 3, pp. 227– 234, Jul. 2005.
- [12] R. Yarlapati, S. Roy Dey, and S. Saha, “Early Prediction of LBW Cases via Minimum Error Rate Classifier: A Statistical Machine Learning Approach,” *2017 IEEE International Conference on Smart Computing (SMARTCOMP)*, May 2017.
- [13] L. C. Y. Poon, N. A. Kametas, N. Maiz, R. Akolekar, and K. H. Nicolaides, “First-Trimester Prediction of Hypertensive Disorders in Pregnancy,” *Hypertension*, vol. 53, no. 5, pp. 812–818, May 2009.
- [14] S. K. Rai and K. N. Sowmya, “A Review on Use of Machine Learning Techniques in Diagnostic Health-Care,” *Artificial Intelligent Systems and Machine Learning*, vol. 10, no. 4, pp. 102– 107, 2018.

- [15] M. Podda, D. Bacciu, A. Micheli, R. Bellù, G. Placidi, and L. Gagliardi, “A machine learning approach to estimating preterm infants survival: development of the Preterm Infants Survival Assessment (PISA) predictor,” *Scientific Reports*, vol. 8, no. 1, Sep. 2018.
- [16] O. Z. Maimon, and L. Rokach, “Data Mining With Decision Trees: Theory And Applications,” vol. 81, 2014.
- [17] L. Rokach and O. Z. Maimon, *Data mining with decision trees: theory and applications*. New Jersey Etc.: World Scientific, Cop, 2015.
- [18] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, “CatBoost: unbiased boosting with categorical features,” *arXiv (Cornell University)*, Jun. 2017.
- [19] A.E Hoerl and R.W. Kennard, “Ridge Regression: Biased Estimation for Nonorthogonal Problems,” *Technometrics*, vol. 42, no.1, pp. 80-86, Feb. 2000.

Final Reports Plagiarism Checking for Mehedi V1

ORIGINALITY REPORT

14%

SIMILARITY INDEX

9%

INTERNET SOURCES

6%

PUBLICATIONS

6%

STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	4%
2	Md Assaduzzaman, Abdullah Al Mamun, Md Zahid Hasan. "Early Prediction of Maternal Health Risk Factors Using Machine Learning Techniques", 2023 International Conference for Advancement in Technology (ICONAT), 2023 Publication	2%
3	dspace.daffodilvarsity.edu.bd:8080 Internet Source	2%
4	journals.plos.org Internet Source	1%
5	Submitted to George Bush High School Student Paper	<1%
6	Submitted to University of Westminster Student Paper	<1%
7	www.mdpi.com Internet Source	<1%