

**Machine Learning-Driven Prediction of Gestational Diabetes Mellitus with WhatsApp Chatbot Integration for Early Risk detection**

A Research Project Submitted to

The **Department of Biomedical Informatics and Biomedical Engineering**

in Partial Fulfilment of the Requirements for the **Degree of Medical Analytics and Informatics** (HMANI)

*by*

**Shamiso Melissa Nyuke**

**R211823M**

Faculty of Medicine and Health Sciences

*© June 2025*

**Supervisor**

**Mr Munyaradzi Dhodho**

SHAMISO M NYUKE RESEARCH PROJECT.docx

University of Zimbabwe



**Document Details**

**Submission ID trn:oid:::6524:290776016**

**Submission Date**

**Jun 23, 2025, 9:32 AM GMT+2**

**Download Date**

**Jun 23, 2025, 9:35 AM GMT+2**

**File Name Shamiso M Nyuke Research Project.docx**

**File Size**

**2.8 MB**

Page 1 of 73 - Cover Page

5% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

**Filtered from the Report**

Bibliography



Quoted Text



Cited Text



Small Matches (less than 10 words)



Submitted works



**Match Groups**

**26**

Not Cited or Quoted 5%



Matches with neither in-text citation nor quotation marks

**0**

Missing Quotations 0%



Matches that are still very similar to source material

**0**

Missing Citation 0%



Matches that have quotation marks, but no in-text citation

**0**

Cited and Quoted 0%



Matches with in-text citation present, but no quotation marks

**Top Sources**

5

%



Internet sources

%

0



Publications

0

%



Submitted works (Student Papers)

**Integrity Flags**

**0 Integrity Flags for Review**

No suspicious text manipulations found.

# DECLARATION

I, **Shamiso Melissa Nyuke, R211823M** declare that the project which is hereby submitted for the qualification of **Bachelor of Science Honours Degree Biomedical Analytics and Informatics** at the **University of Zimbabwe**, is my independent work and has not been handed in before for a qualification at/in another University/Faculty/School.

I further declare that all sources cited or quoted are indicated and acknowledged by means of a comprehensive list of references.

I further cede copyright of the dissertation to University of Zimbabwe.

Signature..................................................................

Student Number R211823M

Date: June 2025

Copyright © 2025 University of Zimbabwe

All rights reserved

# 

# SUPERVISOR APPROVAL FORM

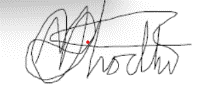
I have received and reviewed the completed research project report to be submitted by the candidate in partial fulfilment of the above degree.

I concur with this submission as final.

Mr Munyaradzi Dhodho (Supervisor), Senior Lecturer,

Department of

Biomedical Informatics and Biomedical Engineering

**Signature…****Date…24/06/25**

# DEDICATION

I would like to dedicate this dissertation to my mother . . .

# ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to Mr Munyaradzi Dhodho, my dissertation supervisor for providing me with the support that was required for this research. I would also like to thank Dr Simbini and the rest of the department for their sincere and valuable guidance and encouragement throughout our academic journey. Finally, I want to thank my mother and my sister Mutsawashe Nyuke for being my support system. And above all, I want to thank the Lord Almighty.

# Abstract

## Introduction:

Gestational Diabetes Mellitus (GDM) is a significant pregnancy complication characterized by glucose intolerance, posing risks to both maternal and fetal health. Early detection is critical to mitigate adverse outcomes, yet many women in low-resource settings, such as Zimbabwe, attend antenatal clinics late, delaying diagnosis. This study explores the integration of machine learning (ML) and a WhatsApp chatbot to provide an accessible, low-cost solution for early GDM risk detection.

## Research Question:

The primary objective was to develop an accurate ML model for predicting GDM risk, while secondary objectives included identifying key maternal risk factors and integrating the model into a WhatsApp chatbot for real-time risk assessment. The central research question focused on the model’s predictive accuracy, with additional questions exploring risk factors and chatbot usability in low-resource settings.

## Methods:

A dataset of 3,525 pregnant women from Kaggle was used to train and evaluate five ML algorithms: Random Forest, Logistic Regression, Support Vector Machine (SVM), Decision Tree, and XGBoost. The dataset included variables such as BMI, OGTT results, family history, prediabetes, gestation in previous pregnancy, blood pressure, lifestyle factors. The best-performing model was integrated into a WhatsApp chatbot using FastAPI, enabling users to input health data and receive instant risk predictions. Model performance was assessed using accuracy, precision, recall, and F1-score, with cross-validation ensuring robustness.

## Results/Findings:

XGBoost emerged as the top-performing model, achieving 96% accuracy, 94% precision, 95% recall, and a 94% F1-score. Key risk factors identified included elevated BMI, abnormal OGTT levels, family history of diabetes, and sedentary lifestyle. The WhatsApp chatbot successfully provided real-time risk assessments during pilot testing, demonstrating its potential as a scalable screening tool. Cross-validation confirmed the model’s reliability, with an average F1-score of 0.897.

## Conclusion:

The study demonstrates that ML models, particularly XGBoost, can effectively predict GDM risk using non-invasive data. The WhatsApp chatbot offers a practical solution for early detection in resource-limited settings, bridging gaps in healthcare access. Future work should focus on clinical validation, expanded datasets, and integration into existing healthcare systems to enhance maternal health outcomes globally. This approach represents a promising step toward equitable and proactive GDM management**.**

Table of Contents

[5% Overall Similarity 2](#_Toc202464257)

[DECLARATION ii](#_Toc202464258)

[SUPERVISOR APPROVAL FORM iii](#_Toc202464259)

[DEDICATION iv](#_Toc202464260)

[ACKNOWLEDGEMENTS v](#_Toc202464261)

[Abstract vi](#_Toc202464262)

[Introduction: vi](#_Toc202464263)

[Research Question: vi](#_Toc202464264)

[Methods: vi](#_Toc202464265)

[Results/Findings: vi](#_Toc202464266)

[Conclusion: vii](#_Toc202464267)

[List Of Abbreviations and acronyms vi](#_Toc202464268)

[Definitions of Key Terms vii](#_Toc202464269)

[CHAPTER 1 : INTRODUCTION 1](#_Toc202464270)

[1.1 Introduction 1](#_Toc202464271)

[1.2 Problem statement : 4](#_Toc202464272)

[1.3 Objectives: 4](#_Toc202464273)

[1.3.1 Primary objective 4](#_Toc202464274)

[1.3.2 Secondary Objective 4](#_Toc202464275)

[1.4 Primary research question 4](#_Toc202464276)

[1.4.1 Secondary research questions 4](#_Toc202464277)

[1.5 Signifance of the study 5](#_Toc202464278)

[1.6 Aim 5](#_Toc202464279)

[1.7 Limitations 5](#_Toc202464280)

[1.8 Conclusion 6](#_Toc202464281)

[CHAPTER 2: LITERATURE REVIEW 7](#_Toc202464282)

[2.1 To predict the risk of gestational diabetes mellitus (GDM) with a machine learning model 7](#_Toc202464283)

[2.2 To identify maternal risk factors contributing to the development of gestational diabetes mellitus 9](#_Toc202464284)

[2.3 To develop a predictive model for GDM machine learning techniques based on the identified risk factors 12](#_Toc202464285)

[2.4 To develop a WhatsApp Chabot where mothers can use to check early warning signs of GDM. 13](#_Toc202464286)

[CHAPTER 3: METHODOLOGY 17](#_Toc202464287)

[3.1 Introduction 17](#_Toc202464288)

[3.2 Study Design 17](#_Toc202464289)

[3.3 Dataset 18](#_Toc202464290)

[3.3.1 Variables 20](#_Toc202464291)

[3.4 Machine Algorithm selection 22](#_Toc202464292)

[3.4.1 Random Forest Model 22](#_Toc202464293)

[3.4.2 Logistic Regression 23](#_Toc202464294)

[3.4.3 Support Vector Machine (SVM) 24](#_Toc202464295)

[3.4.4 Decision Tree 25](#_Toc202464296)

[3.4.5 XGBoost 26](#_Toc202464297)

[3.5 Performance Evaluation 27](#_Toc202464298)

[3.5 Research Materials and Tools 29](#_Toc202464299)

[3.5.1 Software and Frameworks 29](#_Toc202464300)

[3.5.2 Data Visualization: 30](#_Toc202464301)

[3.6 Chatbot Development 30](#_Toc202464302)

[3.6.1 PYTHON FAST API was used to create the backend for the chatbot . 30](#_Toc202464303)

[3.6.2 WHATSAPP API 31](#_Toc202464304)

[3.6.3 Database SQLite was used to create the database to store data from the chatbot . 32](#_Toc202464305)

[3.6.4 Decision Flowchart For the Whatsapp Chatbot 32](#_Toc202464306)

[3.7 Pilot Testing and Validation 34](#_Toc202464307)

[3.8 Ethical Consideration 34](#_Toc202464308)

[3.9 Chapter Summary 34](#_Toc202464309)

[CHAPTER 4: RESULTS AND DISCUSSIONS 35](#_Toc202464310)

[4:1 INTRODUCTION 35](#_Toc202464311)

[4:2 DATA EXPLORATION 35](#_Toc202464312)

[4:3 DATA CLEANING AND EDA 36](#_Toc202464313)

[4.3.1 DATA CLEANING 36](#_Toc202464314)

[4.3.2 Exploratory Data Analysis 38](#_Toc202464315)

[4.3.3 UNIVARIATE ANALYSIS 39](#_Toc202464316)

[4.3.4 Finding outliers 43](#_Toc202464317)

[4.3.5 Outliers handling 45](#_Toc202464318)

[4.3.6 BIVARIATE ANALYSIS 45](#_Toc202464319)

[4. 4 MODEL DEVELOPMENT 47](#_Toc202464320)

[4.4.1 Data Splitting 47](#_Toc202464321)

[4.5 MODEL EVALUATION 48](#_Toc202464322)

[4.5.1 RANDOM FOREST 49](#_Toc202464323)

[4.5.2 LOGISTIC REGRESSION 50](#_Toc202464324)

[4.5.3 DECISION TREE 52](#_Toc202464325)

[4.5.4 SUPPORT VECTOR MACHINE 53](#_Toc202464326)

[4.5.6 XGBoost 55](#_Toc202464327)

[4.6 MODEL COMPARISON 56](#_Toc202464328)

[4.7 MODEL VALIDATION 57](#_Toc202464329)

[4.8 WHATSAPP CHATBOT RESULTS 58](#_Toc202464330)

[4.9 Discussion 61](#_Toc202464331)

[4.9.1 Performance of Machine Learning Models 61](#_Toc202464332)

[4.9.2 Key Risk Factors for GDM 62](#_Toc202464333)

[4.9.3 WhatsApp Chatbot as a Screening Tool 62](#_Toc202464334)

[4.10 Conclusion 63](#_Toc202464335)

[CHAPTER 5: CONCLUSION AND RECOMMENDATIONS 64](#_Toc202464336)

[5.1 Introduction 64](#_Toc202464337)

[5.2 Summary of Findings 64](#_Toc202464338)

[5.2.1.The secondary objectives were also met: 64](#_Toc202464339)

[5.3 Ethical and Practical Considerations 65](#_Toc202464340)

[5.4 Conclusion 65](#_Toc202464341)

[5.5 Recommendations 65](#_Toc202464342)

[REFERENCES 67](#_Toc202464343)

[APPENDICES 71](#_Toc202464344)

[APPENDIX 1: JREC APPROVAL FORM 71](#_Toc202464345)

[APPENDIX 2: Chart Bot Output 72](#_Toc202464346)

**Table of Figures**

[Figure 1 Predictive model flow chart 30](#_Toc201611210)

[Figure 2 Outlook of the variables in the dataset 31](#_Toc201611211)

[Figure 3 Count of patients with GDM and those without GDM 31](#_Toc201611212)

[Figure 4 Random forest model https://medium.com/@roiyeho/random-forests-98892261dc49 35](#_Toc201611213)

[Figure 5 Logistic Regression Algorithm https://www.spiceworks.com/tech/artificial-intelligence/articles/linear-regression-vs-logistic-regression/#lg=1&slide=0 36](#_Toc201611214)

[Figure 6 Support Vector Algorithm https://www.kaggle.com/code/prashant111/svm-classifier-tutorial 37](#_Toc201611215)

[Figure 7 Decision Tree Algorithm https://www.researchgate.net/figure/Decision-tree-structure\_fig1\_360443851 38](#_Toc201611216)

[Figure 8 XGBoost Algorithm https://www.tutorialspoint.com/xgboost/xgboost-architecture.html 39](#_Toc201611217)

[Figure 9 Confusion Matrix ( Draelos 2019) 40](#_Toc201611218)

[Figure 10 Libraries imported 42](#_Toc201611219)

[Figure 11 FastAPI Swagger UI endpoint (predict) 42](#_Toc201611220)

[Figure 12 Endpoint Analysis 43](#_Toc201611221)

[Figure 13 SQLite Database 44](#_Toc201611222)

[Figure 14 Decision Flowchart Whatspp Chatbot 46](#_Toc201611223)

[Figure 15 The dataset contained the variables for the predictions of gestational diabetes mellitus . 47](#_Toc201611224)

[Figure 16 During data exploration the shape of the data had 3525 rows and 17 columns . 48](#_Toc201611225)

[Figure 17 The picture above presents the data types found in the dataset. 48](#_Toc201611226)

[Figure 18 In data cleaning missing values are handled , duplicates removed 49](#_Toc201611227)

[Figure 19 From the picture above it shows no duplicates were found in the dataset 49](#_Toc201611228)

[Figure 20 Missing data was filled using mean because these were continuous variables . 49](#_Toc201611229)

[Figure 21 Summary statistics of the dataset 50](#_Toc201611230)

[Figure 22 Histograms of variables 52](file:///C:\Users\Shamiso\Desktop\DISSERTATION%20PROJECT\SHAMISO%20M%20NYUKE%20RESEARCH%20PROJECT%20(Autosaved).docx#_Toc201611231)

[Figure 23 Histograms of variables 53](file:///C:\Users\Shamiso\Desktop\DISSERTATION%20PROJECT\SHAMISO%20M%20NYUKE%20RESEARCH%20PROJECT%20(Autosaved).docx#_Toc201611232)

[Figure 24 Histograms of variables 54](file:///C:\Users\Shamiso\Desktop\DISSERTATION%20PROJECT\SHAMISO%20M%20NYUKE%20RESEARCH%20PROJECT%20(Autosaved).docx#_Toc201611233)

[Figure 25 Finding Outliers 55](#_Toc201611234)

[Figure 26 Outliers Detected 56](file:///C:\Users\Shamiso\Desktop\DISSERTATION%20PROJECT\SHAMISO%20M%20NYUKE%20RESEARCH%20PROJECT%20(Autosaved).docx#_Toc201611235)

[Figure 27 Outliers Handling 57](#_Toc201611236)

[Figure 28 Correlation Heatmap 58](#_Toc201611237)

[Figure 29 Data Splitting 59](#_Toc201611238)

[Figure 30 Random Forest Evaluation test results 61](#_Toc201611239)

[Figure 31 Random Forest Confusion Matrix 61](#_Toc201611240)

[Figure 32 Random Forest Perfomance Metrics 62](#_Toc201611241)

[Figure 33 Logistic Regression Evaluation Test Results 62](#_Toc201611242)

[Figure 34 Logistic Regression Confusion Matrix 63](#_Toc201611243)

[Figure 35 Logistic Regression Performance Metrics 63](#_Toc201611244)

[Figure 36 Decision Tree Evaluation Test Results 64](#_Toc201611245)

[Figure 37 Decision Tree Confusion Matrix 64](#_Toc201611246)

[Figure 38 Decision Tree Performance Metrics 65](#_Toc201611247)

[Figure 39 Support Vector Machine Evaluation Test Results 65](#_Toc201611248)

[Figure 40 Support Vector Machine Confusion Matrix 66](#_Toc201611249)

[Figure 41 Support Vector Machine Performance Metric Results 66](#_Toc201611250)

[Figure 42 XGBoost Evaluation Test Results 67](#_Toc201611251)

[Figure 43 XGBoost Confusion Matrix 67](#_Toc201611252)

[Figure 44 XGBoost Performance Metrics 68](#_Toc201611253)

[Figure 45 Model Comparison 68](#_Toc201611254)

[Figure 46 Model Validation 69](#_Toc201611255)

[Figure 47 Example of model validation 69](#_Toc201611256)

[Figure 48 example of model validation 70](#_Toc201611257)

[Figure 49 Visual Studio Code executing Pregnancy Pal 70](#_Toc201611258)

[Figure 50 Whatsapp Controller 71](#_Toc201611259)

[Figure 51 Prediction Controller 72](#_Toc201611260)

**Table of tables**

[Table 1 Description Of Variable Statistic 32](#_Toc201611261)

[Table 2 Confusion Matrix Description 41](#_Toc201611262)

[Table 3 Evaluation Metrics Description 42](#_Toc201611263)

# List Of Abbreviations and acronyms

**Medical & Technical Terms:**

1. **GDM –** Gestational Diabetes Mellitus
2. **OGTT –** Oral Glucose Tolerance Test
3. **BMI –** Body Mass Index
4. **HDL –** High-Density Lipoprotein (cholesterol)
5. **PCOS –** Polycystic Ovary Syndrome
6. **Sys BP –** Systolic Blood Pressure
7. **Dia BP –** Diastolic Blood Pressure
8. **T2DM –** Type 2 Diabetes Mellitus
9. **HIP –** Hyperglycemia in Pregnancy
10. **IADPSG**– International Association of Diabetes and Pregnancy Study Groups
11. **NICE –** National Institute for Health and Care Excellence (UK guidelines)

**Machine Learning & Technology:**

1. **ML –** Machine Learning
2. **AI –** Artificial Intelligence
3. **NLP –** Natural Language Processing
4. **XAI –** Explainable Artificial Intelligence
5. **RF –** Random Forest
6. **LR**– Logistic Regression
7. **SVM –** Support Vector Machine
8. **DT –** Decision Tree
9. **MLE –** Maximum Likelihood Estimation
10. **RBF**– Radial Basis Function (kernel in SVM)
11. **API –** Application Programming Interface
12. **IRB –** Institutional Review Board (ethics approval)
13. **SUS –** System Usability Scale

**Organizations & References:**

1. **WHO –** World Health Organization
2. **ADA –** American Diabetes Association
3. **IDF –** International Diabetes Federation
4. **CDC –** Centers for Disease Control and Prevention
5. **ACOG –** American College of Obstetricians and Gynecologists

# Definitions of Key Terms

1. Gestational Diabetes Mellitus (GDM) - A form of diabetes diagnosed during pregnancy, characterized by high blood glucose levels, typically resolving after childbirth but increasing long-term risks of Type 2 diabetes.

2.Oral Glucose Tolerance Test (OGTT) - A diagnostic test for diabetes where blood sugar levels are measured after fasting and again after consuming a glucose-rich drink.

3. Body Mass Index (BMI) - A numerical value calculated from weight and height (kg/m²), used to classify underweight, normal weight, overweight, and obesity.

4. Polycystic Ovary Syndrome (PCOS) - A hormonal disorder in women causing enlarged ovaries with small cysts, often linked to insulin resistance and increased GDM risk.

5. Insulin Resistance- A condition where cells fail to respond effectively to insulin, leading to elevated blood glucose levels—a key factor in GDM.

6. Macrosomia -A birth complication where a newborn weighs >4 kg (8.8 lbs), commonly associated with uncontrolled GDM.

7. FastAPI - A modern Python framework for building APIs, used to connect the ML model to the WhatsApp chatbot.

8. Sedentary Lifestyle -A behavior pattern with minimal physical activity, identified as a modifiable GDM risk factor.

9. Hyperglycemia in Pregnancy (HIP) - An umbrella term for diabetes first detected during pregnancy, including GDM and undiagnosed Type 2 diabetes.

10. System Usability Scale (SUS) - A questionnaire to assess the usability of digital tools (e.g., chatbots).

# CHAPTER 1 : INTRODUCTION

## 1.1 Introduction

During pregnancy, placental hormones can impair insulin function, leading to insulin resistance. While this is typical, excessive resistance may result in gestational diabetes mellitus (GDM). According to International Diabetes Association gestational diabetes (GDM) is caused by high blood glucose that arises during pregnancy and usually disappears after birth but not all times. The mother and baby may experience long-term effects from diabetes during pregnancy. What are the implications? Pregnancy may be affected by complications such as elevated blood pressure, abnormal birth weights (baby girls), and obstructed labour. According to the International Diabetes Federation (IDF) Diabetes Atlas (10th edition, 2021), an estimated 21.1 million live births in 2021 were affected by hyperglycemia in pregnancy (HIP), which includes both gestational diabetes and diabetes first detected during pregnancy. The risk of hyperglycaemia during pregnancy is higher for women over 45, while those with a GDM history are more likely to develop diabetes in their later stages five to ten years after delivery.

Gestational diabetes mellitus (GDM), is one of the most prevalent pregnancy complications which is typified by glucose intolerance initially identified during pregnancy (ADA, 2019). According to (World Health Organisation., 2024) the number of people living with diabetes increased from 200 million in 1990 to 830 million by 2022. Compared to high-income countries, the prevalence rate has been higher in low and middle income. Additionally, more than half of diabetes patients in 2022 did not use medication to manage their condition. Low-and middle-income countries had the lowest level of access to diabetes treatment. Blindness, kidney failure and heart attacks are common complications of diabetes. Over 2 million individuals lost their lives due to diabetes and kidney disease in 2021.Moreover, high blood glucose level was linked to roughly 11% of fatal cardiovascular events. A universal screening of GDM is recommended by WHO during the second trimester (24-28 weeks of gestation) and this would involve an oral glucose tolerance testing (OGTT)which is widely recognized as the standard diagnostic approach.

According to Chen et al., 2015 approximately 7 % of all cases of pregnancy are found to be variedly complicated with GDM and this result in more than 200,000 cases annually. In US only, GDM has been found to complicate about 7-14 % cases annually, and the trend seems to have increased by 35-100 % in the recent years. A history of GDM can be considered to be one of the sturdiest risk factors concerning the development of type 2 diabetes. Among women who have a history of GDM, the risk of developing classical type 2 diabetes usually ranges from 20 to 50 %. Evidences collected from various efficacy trials suggest that lifestyle interventions like weight management can modulate and prevent type 2 diabetes in at-risk individuals. The cornerstone of GDM management is glycemic control, and hence, it is attributed to be the main focus of attention for the therapy. In this review, we have tried to highlight the various risk factors associated with GDM along with the available therapeutic options in the treatment and management of the disease.

According to the (Kansu-Celik et al.,2019), GDM is a form of glucose intolerance that was first recognized in pregnancy. It is estimated to affect approximately 1 in 7 births globally, with rates varying from just 1% across different regions to more than 30% in some populations. Gestational Diabetes Mellitus complicates as high as 14% of pregnancies in the United States, with an annual count of 20000000 cases. It is linked to changing lifestyles, higher maternal age, or obesity levels. GDM is believed to have a global prevalence of approximately 16%, with South-East Asia and the Middle East having one of the highest rates. As an illustration, GDM is present in as much as 27% of pregnant women in urban areas across India. It is known to cause adverse maternal outcomes such as pre-eclampsia, caesarean delivery and long term risks of type 2 diabetes for mothers. New-born complications that may arise include macrosomia, hypoglycaemia, and an increased risk of respiratory distress. Between 8% and 18% of the population in Canada is GDM, while in China it ranges from 6.8% to 10.4% (Hirst et al., 2012). India has an estimated 27.5% of the population estimated to be GDM-positive, whereas Sri Lanka and Bangladesh have only 9.9% and 9.8% of this category. New screening criteria for GDM were recently recommended by the International Association of Diabetes and Pregnancy Study Groups (IADPSG) in response to the HAPO study. Using these criteria, the total incidence of GDM is nearly 15%-20% (ADA., 2014).

In North America, GDM is prevalent in a range of 5 to 10% according to CDC meta-analysis in 2023, with the majority of cases attributed to GDF and others to pre-existing diabetes. Those in Asia exhibit higher rates, with India reporting 10-14% of cases, and urban areas experiencing up to 27%. The Middle East may witness a rise in obesity due to lifestyle changes, with rates exceeding 20%. The prevalence of GDM in sub-Saharan Africa varies between 8% and 14%, according to different studies.

Sub-Saharan Africa is experiencing a growing problem of gestational diabetes mellitus (GDM), both locally and globally. According to research, the prevalence of GDM in sub-Saharan Africa varies significantly depending on demographic factors and diagnostic criteria. In Cameroon, GDM prevalence varies by diagnostic criteria: 5.9% (WHO), 17.7% (IADPSG), and 11.0% (NICE guidelines). In Uganda, antenatal clinic studies have revealed a GDM prevalence of 6.4%, which highlights risk factors such as higher maternal age, obesity and aging mothers, among others. However, only women aged 30 to 49 were reported to have met this recommendation by 75% in older women. This occurrence corresponds with other local findings, which highlight the impact of lifestyle modifications, urbanization, and limited diagnostic resources.GDM is still underdiagnosed in many low- and middle income countries, including those in sub-Saharan Africa, despite these figures. The use of inaccurate urine glucose tests remains widespread in certain nations, resulting in missed diagnoses and delayed interventions.

Unfortunately, babies born to mothers with GDM are more likely to be too big for their gestational age, and women with GDM may develop Type 2 diabetes (T2DM) later in life. A Zimbabwean study (Nhidza et al., 2018) screened 150 pregnant women at Parirenyatwa Antenatal Clinic using WHO criteria. Of these, 17 underwent OGTT testing, revealing a 6.7% GDM prevalence (10 cases)—comparable to Indian studies but lower than other African research, likely due to differing diagnostic approaches. All pregnant women who gave their consent had urine samples taken. A random blood sugar analysis was then performed if a woman reported clinical symptoms of GDM and a urinalysis revealed glycosuria. Following the collection of a fasting blood sugar sample, those who were suspected of having GDM due to elevated glucose (n=17) were screened with the glucose load challenge the next day. Self-reported family history of diabetes. Consenting women (N=150) between 24 and 28 weeks of gestation were enlisted. About half of the participants were gradiva 1, and their mean age was 27.2 (3.5) years. While none of the participants had a maternal history of type 2 diabetes, they did report other family members who did. In conclusion, the prevalence of GDM is comparable to one Indian study but lower than two comparable African studies. It is noteworthy that different diagnostic criteria may be the cause of reported prevalence differences across populations from various studies ( Nhidza et al., 2018)

## 1.2 Problem statement :

Timely detection of gestational diabetes mellitus (GDM) is critical for preventing complications during pregnancy. However, during my visits at one of the hospitals in Harare I discovered that a significant number of pregnant women attend antenatal clinics late, often just before delivery. This limits the time available for adequate screening and early intervention, such as for conditions like GDM. The lack of early detection not only increases the risk to maternal and fetal health but also places a strain on healthcare providers. There is a pressing need for accessible, early risk detection tools to address this gap.

## 1.3 Objectives:

### 1.3.1 Primary objective

To predict the risk of gestational diabetes mellitus (GDM) with a machine learning model

### 1.3.2 Secondary Objective

1. To identify maternal risk factors contributing to the development of gestational diabetes mellitus

2. To develop a predictive model for GDM machine learning techniques based on the identified risk factors

3.To develop a WhatsApp chatbot where mothers can use to check early warning signs of GDM

## 1.4 Primary research question

How accurately can a machine learning model predict the risk of gestational diabetes mellitus (GDM) in pregnant women.

### 1.4.1 Secondary research questions

1. What are the maternal risk factors that contribute to the development of gestational diabetes mellitus (GDM)?

2. How can a predictive model for GDM be developed machine learning techniques based on identified risk factors?

3. How can a whatsapp Chabot be developed to support early diagnosis and management of GDM in low-resource settings?

## 1.5 Signifance of the study

Gestational Diabetes Mellitus (GDM) is a growing public health concern in Zimbabwe, with significant risks to maternal and fetal health. While hypertension has been identified as a leading cause of maternal mortality in Zimbabwe (HealthTimes, 2025), GDM remains an underdiagnosed yet equally dangerous complication. The article highlights that hypertensive disorders in pregnancy contribute to severe outcomes, but GDM—often coexisting with or exacerbating hypertension—also leads to adverse effects such as macrosomia, preterm birth, and long-term metabolic disorders for both mother and child.

This study addresses a critical gap in early GDM detection by developing a machine learning (ML)-powered WhatsApp chatbot that provides accessible, low-cost risk assessment for pregnant women, particularly in resource-limited settings like Zimbabwe. Many women in Harare attend antenatal clinics late, often just before delivery, delaying timely screening and intervention for GDM.

The proposed solution leverages ML algorithms to predict GDM risk using easily obtainable clinical and demographic data (e.g., BMI, age, family history) and integrates this model into a WhatsApp chatbot—a platform widely used in Zimbabwe. This approach aligns with the need for innovative, scalable healthcare tools in low-income regions, as emphasized by the HealthTimes (2025) report on maternal health challenges.

## 1.6 Aim

The aim of this study is to develop an accessible and accurate machine learning (ML)-based predictive model for early detection of Gestational Diabetes Mellitus (GDM), integrated with a WhatsApp chatbot, to improve screening and risk assessment in low-resource settings.

## 1.7 Limitations

1.Using secondary data (since no collection of primary data from patients was done).

2.A lack of clinical evaluation

3.Since the dataset is limited, rely on factors that don’t need invasive data, for example, age and health records.

## 1.8 Conclusion

Gestational diabetes mellitus (GDM) is associated with considerable threats to both maternal and foetal lives, and its prevalence is on the increase world over as well as in sub-Saharan Africa within Zimbabwe. Although GDM has been at times fatal, its diagnosis has been underreported because women get their antenatal check later and have more limited screening options in low-income areas. This research aims at filling this gap by proposing a machine learning-powered predictive model that would be implemented as a WhatsApp chatbot, a tool that would be both affordable and convenient to use in assessing the risk of developing GDM at a very early stage. This study will provide a performance enhancement of the early identification and management of underlying risks, as it will utilize understandable, accessible technology with the easing of non-invasive risk factors that will reduce adverse pregnancy outcomes linked to cases of GDM. The approach will be discussed in terms of its methodology, implementation, and evaluation in further chapters.

# 

# CHAPTER 2: LITERATURE REVIEW

## 2.1 To predict the risk of gestational diabetes mellitus (GDM) with a machine learning model

Gestational Diabetes Mellitus (GDM) is a prevalent complication during pregnancy, characterised by high blood sugar levels that are first recognised during pregnancy (ADA 2023). Early detection and management of GDM are critical to prevent adverse maternal and fetal outcomes, such as macrosomia, preterm delivery, and preeclampsia (McIntyre et al. 2022). Traditional gestational diabetes screening methods like the oral glucose tolerance test (OGTT) present notable practical challenges, requiring fasting protocols, multiple blood samples and several healthcare visits - particular obstacles in resource-constrained environments (Sacks et al. 2022; Zhou et al. 2021). Machine learning approaches have emerged as a promising alternative, with studies demonstrating 85-95% prediction accuracy using routinely available clinical and demographic data (Artzi et al. 2020; Li et al. 2023). The World Health Organization has specifically highlighted such digital tools as viable solutions for settings with limited diagnostic infrastructure (WHO 2021).

Machine learning (ML) models, particularly ensemble methods like XGBoost and Random Forest, have shown promise in GDM prediction by leveraging clinical and demographic data. For example, Zaky et al. (2025) trained an ensemble model on first-trimester biomarkers (e.g., glucose levels, insulin resistance), achieving 89% accuracy. However, their reliance on blood-based biomarkers limits scalability in resource-limited settings."The model highlighted the importance of biomarkers such as a history of high glucose levels, insulin, and cholesterol as key predictive features. Similarly, Santhadevi et al. (2025) evaluated multiple ML and deep learning models, with XGBoost and Random Forest achieving accuracies of 95.75% and 90.5%, respectively. These studies underscore the effectiveness of ensemble methods in minimizing overfitting and enhancing predictive power. However, both studies relied on biomarkers that require blood sample collection, which may not be feasible in low-resource settings.

Hassan et al. (2025) advanced the field further by proposing a fusion model that combined the strengths of ML and deep learning algorithms. Their model achieved a high accuracy of 98.21%, with interpretable features identified using Explainable Artificial Intelligence (XAI) techniques. The inclusion of feature engineering and oversampling techniques to address class imbalance was a notable strength of this study. However, the reliance on complex models may pose challenges for real-world implementation, particularly in regions with limited computational resources.

In contrast, Damayanti and Baita (2025) focused on simpler ML models, such as Support Vector Machines (SVM) and Random Forest (RF), achieving 100% and 99% accuracy, respectively, after applying random undersampling to balance the target data. Their findings suggest that SVM may outperform RF in terms of misprediction rates, offering a more reliable option for GDM prediction. On the other hand, the small dataset used in this study raises concerns about the generalizability of the results.

Kaya et al. (2024) adopted a different approach by focusing on sociodemographic and obstetric features, such as maternal age, body mass index (BMI), and family history of diabetes, to predict GDM. Their study highlighted the importance of venous plasma glucose levels and BMI as key predictors, with the XGB Classifier achieving an accuracy of 72.7%. While this approach reduces the reliance on biomarkers, the relatively lower accuracy compared to other studies suggests that further refinement of feature selection and model optimization is needed.

Despite the progress made in GDM prediction using ML models, several limitations remain. Many studies, such as those by Zaky et al. (2025) and Kaya et al. (2024), were conducted on relatively small datasets, which may not fully capture the diversity of GDM cases. Additionally, the dependence on biomarkers requiring blood sample collection limits the applicability of these models in resource-constrained settings. Furthermore, while some studies have proposed hybrid models and advanced techniques, the clinical implementation of these models remains a challenge, particularly in low-income countries.

The proposed project addresses several gaps identified in the existing literature. While previous studies have demonstrated the efficacy of ML models for GDM prediction, few have focused on developing user-friendly tools that can be integrated into routine clinical practice, particularly in low-resource settings. This project builds upon the strengths of existing ML models, such as those developed by Zaky et al. (2025) and Hassan et al. (2025), by incorporating a WhatsApp chatbot for real-time risk assessment and early detection. The integration of a WhatsApp chatbot offers several advantages.

Firstly, it provides a low-cost and widely accessible platform for GDM screening, making it suitable for resource-limited healthcare systems. Secondly, the chatbot enables remote monitoring of health indicators, allowing healthcare providers to send timely alerts and recommendations to at-risk patients. This approach aligns with the recommendations of Hassan et al. (2025) for the development of mobile and web applications to facilitate proactive GDM management. Moreover, the project differs from previous studies by focusing on early risk detection through a user-friendly interface, which has not been explored in depth in the existing literature. While studies such as those by Damayanti and Baita (2025) have achieved high accuracy with ML models, they have not addressed the practical challenges of implementing these models in real-world clinical settings. The proposed project bridges this gap by developing a tool that is both accurate and accessible, ensuring that GDM prediction is not only reliable but also practical for widespread use. In conclusion, this project advances the field of GDM prediction by combining the strengths of ML models with the accessibility of a WhatsApp chatbot. By addressing the limitations of existing studies, such as dataset size, biomarker dependency, and clinical implementation, this project offers a novel solution for early GDM detection that is both effective and scalable.

## 2.2 To identify maternal risk factors contributing to the development of gestational diabetes mellitus

The identification of maternal risk factors for gestational diabetes mellitus (GDM) forms the foundation for effective screening protocols and preventive interventions (ACOG 2023). Timely diagnosis and treatment of GDM are proven to mitigate adverse outcomes including fetal macrosomia and hypertensive disorders (Metzger et al. 2008; McIntyre et al. 2022), necessitating their prioritisation in antenatal care frameworks (WHO 2023).This review aims to synthesize existing literature on the risk factors contributing to GDM and highlight the gaps that the proposed project seeks to address. Associated with adverse maternal and fetal outcomes, including increased risks of cesarean delivery, preterm birth, and neonatal complications (Imam et al., 2025; Fortofoiu et al., 2022). Several studies have identified modifiable risk factors that contribute to the development of GDM. Modifiable risk factors like obesity (adjusted OR: 1.63; Mirabelli et al., 2023) and hypertension dominate GDM development, while non-modifiable factors (e.g., age >35, prior macrosomia) further stratify risk. Our model will weight these features using SHAP values to improve interpretability.

Imam et al. (2025) conducted a study involving 200 Romanian pregnant women and found that overweight and hypertensive women had a higher prevalence of GDM. Specifically, 25% of overweight women and 12% of hypertensive women developed GDM, with a significant association between urban residence and these risk factors. These findings align with those of Cascabulho et al. (2024), who emphasized the role of lifestyle factors, including poor eating habits and obesity, in the development of GDM. Both studies underscore the importance of addressing modifiable lifestyle and environmental factors in GDM prevention. Mirabelli et al. (2023) further expanded on the role of modifiable risk factors, highlighting maternal preconception body mass index (BMI) as a significant predictor of GDM. In a study of 3,856 singleton pregnancies, the authors found that obese women had a 2.525-fold increased risk of developing GDM compared to those with a normal BMI. This study also noted that being overweight (BMI 25–29.9) independently increased the risk of GDM, with an adjusted odds ratio (OR) of 1.63. These findings suggest that preconception weight management could play a critical role in reducing GDM incidence.

In addition to modifiable factors, several non-modifiable risk factors have been identified. Advanced maternal age, defined as 35 years or older, has been consistently associated with an increased risk of GDM (Mirabelli et al., 2023). Ethnicity also plays a role, with certain ethnic groups demonstrating a higher predisposition to GDM. Furthermore, a history of spontaneous abortions and previous GDM diagnosis are non-modifiable risk factors that significantly increase the likelihood of developing GDM in subsequent pregnancies. Fortofoiu et al. (2022) identified additional non-modifiable risk factors, including a history of fetal macrosomia (birth of large babies) and pre-existing hypertension. In their study of 97 pregnant women, 13.7% of those with GDM had a history of delivering large babies, compared to none in the non-GDM group. This finding highlights the potential role of fetal macrosomia as a previously underrecognized risk factor for GDM. The study also noted that 7.8% of women with GDM had a history of GDM in previous pregnancies, further emphasizing the importance of obstetric history in GDM risk assessment.

The development of GDM is also influenced by genetic and environmental factors. Rosik et al. (2020) highlighted the role of genetic variants, such as KCNJ11, GCK, and HNF1A, which are associated with insulin secretion and glucose metabolism. These genetic factors, identified through genome-wide association studies, provide insights into the molecular mechanisms underlying GDM. However, the interplay between genetic predisposition and environmental factors, such as diet and lifestyle, remains poorly understood. This multifactorial nature of GDM underscores the need for a holistic approach to risk factor identification and management.

While the studies reviewed have significantly advanced our understanding of GDM risk factors, several gaps remain. Fortofoiu et al. (2022) noted that their study was limited by a small sample size of 97 women, which may restrict the generalizability of their findings. Similarly, the study's cross-sectional design, which assessed risk factors at a single time point (24–28 weeks of pregnancy), limits the ability to track changes in risk factors over time. Longitudinal studies would provide a more comprehensive understanding of how risk factors evolve during pregnancy and their impact on GDM development. Additionally, existing studies have primarily focused on clinical and demographic risk factors, with limited exploration of how these factors can be integrated into early detection tools. The use of machine learning models to predict GDM based on multiple risk factors has not been widely explored. Furthermore, the integration of such models into user-friendly platforms, such as WhatsApp chatbots, has not been addressed in the existing literature. These gaps highlight the need for innovative approaches to improve early risk detection and management of GDM.

The proposed project addresses several gaps in the existing literature. First, the integration of machine learning models allows for the simultaneous consideration of multiple risk factors, including both modifiable and non-modifiable factors, to predict GDM with higher accuracy. This approach builds on the findings of Imam et al. (2025), Mirabelli et al. (2023), and Fortofoiu et al. (2022), who identified key risk factors but did not explore their combined predictive potential. Second, the use of a WhatsApp chatbot represents a novel approach to early risk detection. WhatsApp is widely used globally, making it an accessible platform for pregnant women, particularly in low-resource settings. This approach addresses the need for early and equitable access to GDM screening, as highlighted by the limitations of existing studies in terms of generalizability and accessibility. In conclusion, the proposed project offers a unique and innovative solution to the early detection of GDM by leveraging machine learning models and a widely accessible communication platform. By addressing the gaps in existing literature, this project has the potential to improve maternal and fetal outcomes for women at risk of GDM.

## 2.3 To develop a predictive model for GDM machine learning techniques based on the identified risk factors

Advancements in machine learning (ML) have recently provided promising avenues for predicting GDM, leveraging various risk factors and datasets. Several studies have demonstrated the efficacy of ML models in predicting GDM, often outperforming traditional statistical methods. Belsti et al. (2023) reported that ML approaches achieved superior predictive performance, with accuracies ranging from 75% to 93%, compared to conventional methods. Their study utilized routine antenatal care data, emphasizing the importance of established risk factors such as GDM history, BMI, and ethnic descent. The CatBoost Classifier emerged as the top performer, achieving an accuracy of 85%, precision of 90%, and recall of 78%. Similarly, Tiwari et al. (2024) explored ML classifiers, achieving accuracies between 90.5% and 95%, with KNN and XGBoost demonstrating the best results. These findings underscore the potential of ML in developing efficient GDM diagnosis systems, particularly in regions with limited diagnostic services.

Bigdeli et al. (2025) further advanced this field by developing six ML models to predict GDM in the first trimester, with Random Forest (RF) showing the best performance (Accuracy: 89%, Precision: 86%, Recall: 92%, AUC: 94%). The study highlighted the importance of early diagnosis, allowing timely interventions to improve maternal and fetal health outcomes. These studies collectively demonstrate the effectiveness of ML in GDM prediction, particularly when leveraging established risk factors and clinical parameters.

Wang et al. (2022) and Zhang & Wang (2022) introduced ensemble methods to enhance GDM prediction accuracy. Wang et al. proposed an integrated LightGBM-XGBoost-GB model, achieving a 2.56% increase in accuracy compared to standalone models. Their study emphasized the importance of clinical risk factors such as age, BMI, and genetic indicators. Similarly, Zhang & Wang utilized ensemble learning with algorithms like XGBoost and LightGBM, achieving 80.3% accuracy. Both studies highlighted feature engineering and selection as critical for model accuracy, particularly in small datasets.

However, these studies had notable limitations. While Wang et al. (2022) achieved high accuracy (92%) with an ensemble model, the absence of external validation on diverse cohorts—particularly from low-income regions—limits clinical applicability. This gap underscores the need for cross-population testing in our proposed model. Zhang & Wang did not discuss clinical validation, limiting the assessment of real-world applicability. These gaps underscore the need for models that are not only accurate but also validated across diverse populations and clinical settings.

The current project aims to develop a predictive model for GDM using ML techniques, integrated with a WhatsApp chatbot for early risk detection. This approach addresses several gaps identified in the literature. Firstly, while existing studies have focused on model development, few have explored the practical application of these models in real-world settings. The integration with a WhatsApp chatbot provides a user-friendly platform for early risk assessment, making it accessible for widespread use, particularly in regions with limited healthcare resources.

Secondly, the project will incorporate a broader range of risk factors, including genetic predispositions, lifestyle, and environmental influences, which were not comprehensively addressed in previous studies. This comprehensive approach aims to enhance model robustness and accuracy.

Finally, the project will emphasize external validation to ensure the model's applicability across diverse populations, addressing a critical limitation of earlier studies. By leveraging ensemble methods and feature engineering, the project seeks to optimize model performance and reliability.

In conclusion, while previous studies have made significant contributions to GDM prediction using ML, the current project offers a novel approach by integrating predictive models with a WhatsApp chatbot, addressing gaps in risk factor consideration, and ensuring external validation. This innovative strategy has the potential to enhance early detection and improve health outcomes for pregnant women globally.

## 2.4 To develop a WhatsApp Chabot where mothers can use to check early warning signs of GDM.

Gestational Diabetes Mellitus (GDM) is a significant health concern during pregnancy, often leading to adverse maternal and fetal outcomes if left undiagnosed or poorly managed. Early detection and timely intervention are critical to improving health outcomes for both mothers and babies. Traditional methods of GDM screening, such as glucose tolerance tests, are effective but may not be readily accessible to all populations, particularly in rural or resource-limited settings. The integration of digital health technologies, such as machine learning models and chatbots, offers a promising solution to address these challenges by enabling early risk detection and providing personalized health advice.

The use of chatbots in healthcare has gained traction in recent years due to their ability to provide instant, accessible, and personalized health information. These platforms leverage machine learning (ML) and natural language processing (NLP) to deliver accurate and context-specific responses, making them particularly useful for maternal health support. This literature review explores the existing body of work on chatbots for maternal health, with a focus on their application in GDM prediction and early risk detection.

Mugoye et al. (2019) emphasized the importance of providing immediate and accurate health information to pregnant women through digital platforms. They highlighted the limitations of existing health information systems, particularly in rural areas, where access to healthcare providers is limited. The authors proposed the development of mobile phone-integrated chatbots to address the information gaps faced by expectant mothers. These chatbots would utilize AI and machine learning algorithms to deliver precise and timely responses to pregnancy-related queries.

Puspitasari et al. (2022) built on this idea by developing a semi-automated chatbot for maternal health education and monitoring. While their study did not specifically focus on GDM, it underscored the importance of providing immediate responses to physical complaints and danger signs during pregnancy. The chatbot was designed to deliver quick and accurate information, which is critical for early detection of complications and timely consultation with healthcare providers. The study identified three major themes: maternal health education, information on maternal health services, and health monitoring. These themes guided the development of the chatbot, ensuring that it addressed the diverse needs of pregnant women.Both studies agree on the potential of chatbots to improve maternal health outcomes by providing accessible and timely health information. However, they also highlight the need for further research to enhance the accuracy and usability of these platforms.

Bhaskar et al. (2024) introduced PregBot, a chatbot system that leverages machine learning and NLP to provide personalized maternal health support. The system was designed to offer customized guidelines and real-time query resolution, addressing the unique needs of pregnant women at different stages of their pregnancy. PregBot also incorporated a community-building feature, enabling users to share experiences and advice, which further enhanced the overall pregnancy journey. The study demonstrated the transformative potential of ML and NLP in delivering precise and context-specific health information.

Similarly, Sagstad et al. (2022) developed Dina, a chatbot specifically designed for GDM patients. The chatbot was integrated into Norway's national digital health platform and focused on addressing questions related to blood glucose levels, diet, and physical activity. The study found that the chatbot successfully processed and responded to 88.51% of user queries, highlighting its effectiveness as a tool for GDM management. However, the study also identified several limitations, including the chatbot's availability only in Norwegian and the lack of a feedback mechanism to assess user satisfaction.

Afrizal et al. (2022) developed a Telegram-based chatbot for maternal health monitoring, which utilized NLP to recognize user inputs and provide personalized health advice. The chatbot was designed to educate users about danger signs during pregnancy and facilitate remote counseling. The study reported a System Usability Scale (SUS) score of 62.3, indicating that the chatbot was usable but had room for improvement in terms of user experience.

These studies demonstrate the growing interest in leveraging ML and NLP to develop chatbots for maternal health. However, they also highlight the need for further refinement to address issues such as language barriers, usability, and clinical validation.

While the studies discussed above have made significant contributions to the field of maternal health chatbots, several gaps remain. Mugoye et al. (2019) and Puspitasari et al. (2022) focused primarily on the development of chatbots for general maternal health education and monitoring, with limited emphasis on GDM prediction. Bhaskar et al. (2024) and Sagstad et al. (2022) addressed GDM more specifically but faced limitations such as language barriers and the lack of feedback mechanisms. Afrizal et al. (2022) highlighted the importance of usability and clinical validation, which were not adequately addressed in their study.

Furthermore, none of the studies integrated machine learning models for GDM prediction into their chatbots. This represents a significant gap, as early detection of GDM is critical for improving health outcomes. Additionally, the studies relied heavily on qualitative methods, which limited the generalizability of their findings. There is a need for studies that combine qualitative and quantitative approaches to evaluate the effectiveness of chatbots more comprehensively.

The proposed project addresses the gaps identified in the literature. It fills gaps in detection gestational diabetes by offering a smart WhatsApp chatbot that uses machine learning together with a platform most individuals can easily use in these areas. It uses basic facts about the person’s health checked through text chat (BMI, family notes) and predicts their disease risk, providing results quickly and using fewer clinic visits. Because the chatbot is easy on bandwidth and easy to expand, people get advised by it and can offer input to keep improving it. Using both numbers and user opinions together in this approach helps overcome what prior studies lacked and still fulfills WHO criteria for digital health apps in low-resource groups.

In conclusion, the proposed project builds on the existing body of work on maternal health chatbots while addressing the gaps identified in the literature. By integrating machine learning models with a WhatsApp chatbot, the project offers a novel and effective solution for GDM prediction and early risk detection. The project has the potential to improve maternal health outcomes globally, particularly in resource-limited settings.

# CHAPTER 3: METHODOLOGY

## 3.1 Introduction

This chapter outlines the methodology employed in the study to develop a machine learning (ML)-driven predictive model for Gestational Diabetes Mellitus (GDM) risk assessment, integrated with a WhatsApp chatbot for early detection. The chapter details the research design, dataset, variables, operationalisation, data collection, validity and reliability measures, data analysis procedures, and ethical considerations. The methodology ensures a systematic, transparent, and reproducible approach to achieving the study’s objectives.

## 3.2 Study Design

This study aims to develop a predictive machine learning models capable of predicting gestational diabetes mellitus and detect early risk from a whatsapp chatbot . Five algorithms Support Vector Machine (SVM), Decision Tree, Logistic Regression, Random Forest, and XGBoost - will be implemented and compared for their diagnostic accuracy. The research will progress through four critical phases: (1) data collection and preprocessing of clinical symptom datasets, (2) feature selection and engineering to identify key diagnostic markers, (3) model training and optimization using cross-validation techniques, and (4) comprehensive performance evaluation to determine the most effective classifier. The subsequent sections detail each methodological stage of this predictive modeling approach.

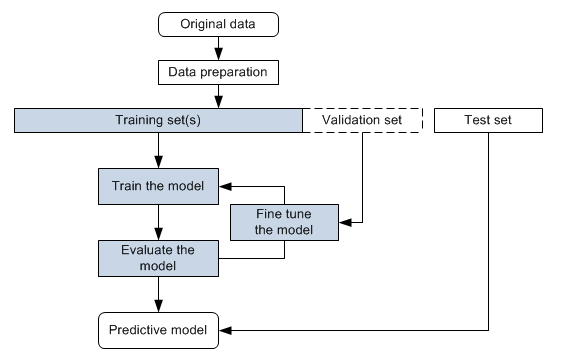


Figure 1 Predictive model flow chart

## 3.3 Dataset

This study used data that comes from Kaggle which is commonly used by many in the world of data science and machine learning. Including data on BMI, family background, OGTT findings and lifestyle choices, the set of records encompasses the medical and demographic records of 3,525 pregnant women and includes variables vital for predicting GDM. The data was selected since it is broad and addresses the same risk factors used in the definition of GDM, so the model would be reliable. The data was separated into respective training, validation and testing sets to allow for smooth model building, configuration and proper testing (Kaggle, 2024).

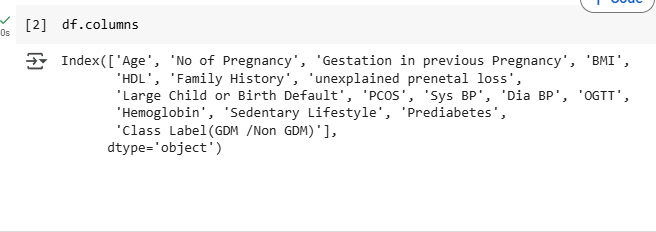


Figure 2 Outlook of the variables in the dataset



Figure 3 Count of patients with GDM and those without GDM

The above figure was the count for the dataset. It showed class imbalance as Non GDM instances were higher than GDM instances . This imbalance could cause bias machine learning models toward the majority class, leading to poor performance in predicting the minority class (GDM cases), which is clinically critical for early diagnosis (Chawla et al., 2002). To address this, class weights were computed and applied during model training (e.g., in Random Forest and Logistic Regression) to penalize misclassification of the minority class more heavily. This technique adjusts the loss function to prioritize GDM cases, improving recall without sacrificing precision (Fernández et al., 2018). For example, assigning higher weights to GDM samples forces the model to learn their patterns more effectively. Studies show that weighting mitigates bias in imbalanced medical datasets, ensuring clinically actionable predictions (Johnson & Khoshgoftaar, 2019).

## 

## 3.3.1 Variables

Table 1 Description Of Variable Statistic

|  |  |  |
| --- | --- | --- |
| Variable | Description | Significance |
| Age | The age of the individual in years. | Older maternal age is associated with a higher risk of gestational diabetes. For example, women aged 35 and older often face increased metabolic challenges during pregnancy. |
| No of Pregnancy | The total number of pregnancies the individual has had. (n= 1,2,3,4) | A higher number of pregnancies can indicate a greater likelihood of developing gestational diabetes due to cumulative physiological stress and potential complications. |
| Gestation in Previous Pregnancy | Refers to the length of time a woman was pregnant and measured in weeks. Infant born before 37wks is premature and one who is born after 42wks is post mature .  Normal range =37wks -42wks  In the dataset 1indicates <37wks , 0 indicate =37 wks-42wks, 2 indicates >42wks | Preterm birth and post-term pregnancy increases the chance of complications like gestational diabetes mellitus . |
| BMI (Body Mass Index) | A measure of body fat based on height and weight, calculated as weight in kilograms divided by height in meters squared (kg/m²). Normal range is from 18.5-24.9. Underweight <18.5. Overweight =25-29.9. Obese >= 30. | A BMI of 30 or higher categorizes an individual as obese, which is a known risk factor for developing gestational diabetes. For instance, a BMI of 55 indicates severe obesity. |
| HDL (High-Density Lipoprotein) | The level of HDL cholesterol in mg/dL. Normal range is >=50mg/dl. | Higher levels of HDL cholesterol are generally protective against cardiovascular disease. Low levels (<40mg/dl) may indicate metabolic disturbances that can increase the risk of diabetes. |
| Family History | Indicates whether there is a family history of diabetes (1 = Yes, 0 = No). | A family history of diabetes increases the risk of gestational diabetes if a first degree relative has diabetes , due to genetic and environmental factors. This variable helps assess predisposition. |
| Unexplained Prenatal Loss | Indicates whether there has been any unexplained loss in previous pregnancies (1 = Yes, 0 = No). | This factor can indicate underlying health issues that may affect future pregnancies, including the risk of diabetes. Loss of pregnancy without a known cause or explanation. |
| Large Child or Birth Defect | Indicates whether the individual has had a large child (macrosomia) or any birth defects (1 = Yes, 0 = No). | Having previously given birth to a large child is a strong indicator of insulin resistance and can signal a higher risk for developing gestational diabetes. |
| PCOS (Polycystic Ovary Syndrome) | Shows whether the individual has been diagnosed with PCOS (1 = Yes, 0 = No). | PCOS is associated with insulin resistance and an increased risk of developing diabetes, including gestational diabetes during pregnancy. |
| Sys BP (Systolic Blood Pressure) | Indicates systolic blood pressure in mmHg.Normal <120mmHg. | Elevated blood pressure can be a sign of gestational hypertension or preeclampsia, which are conditions related to increased risk for diabetes. |
| Dia BP (Diastolic Blood Pressure) | Indicates diastolic blood pressure in mmHg.Normal<80mmHg. | Similar to systolic blood pressure, high diastolic values may indicate cardiovascular issues that could affect diabetes risk during pregnancy. |
| OGTT (Oral Glucose Tolerance Test) | Represents the results of the OGTT in mg/dL, which is crucial for diagnosing diabetes. Normal range for pregnant women is (1hr-<140mg/dl)  3hr , fasting<95mg/dl, 1-hour<180mg/dl,2-hour<155mg/dl,3-hour<140mg/dl. Higher than 200mg/dl or higher is at a risk . | Abnormal OGTT levels are used to diagnose gestational diabetes. Values above certain thresholds indicate impaired glucose metabolism. |
| Hemoglobin | Indicates hemoglobin levels in g/dL. Ranges from 11-14 for pregnant women . | Hemoglobin levels can indicate overall health and nutritional status. Low levels may suggest anemia, which can complicate pregnancy management. |
| Sedentary Lifestyle | Indicates whether the individual leads a sedentary lifestyle (1 = Yes, 0 = No). | A sedentary lifestyle is associated with higher risks of obesity and diabetes, including gestational diabetes. Lifestyle interventions can be critical in risk management. |
| Prediabetes | Indicates whether the individual is prediabetic (1 = Yes, 0 = No). | Prediabetes is a significant risk factor for developing gestational diabetes, as it indicates existing glucose intolerance. |
| Class Label (GDM / Non GDM) | The target variable indicating whether the individual has gestational diabetes mellitus (1 = GDM, 0 = Non-GDM). | This is the outcome variable in the dataset. It is essential for evaluating the effectiveness of risk factors and developing predictive models for gestational diabetes. |

## 3.4 Machine Algorithm selection

The dataset was split into training (70%), validation (15%), and test (15%) subsets to ensure robust model development, optimization, and unbiased evaluation. The approach taken in this study was to select appropriate machine learning models for predicting Gestational Diabetes Mellitus (GDM). The objective was to design models that do well in diagnosis and also explain their decisions and manage data that is not evenly distributed. Random Forest (RF), Logistic Regression (LR), Support Vector Machine (SVM), Decision Tree (DT) and XGBoost were used in this evaluation.

### 3.4.1 Random Forest Model

A random forest takes several decision trees and forms the output by taking the mode (for classification) or mean (for regression) of their predictions which increases accuracy and reduces the risk of overfitting (Breiman, 2001). In this process, each tree uses a random part of the whole dataset and it chooses different features randomly to split the tree, making the trees diverse. By bagging several trees, the influence of each tree’s mistakes is reduced, helping the model become robust and used in many situations. Random forests are used a lot since they do well, scale well and can manage large data sets (Liaw & Wiener, 2002). New studies have proven that it can accurately estimate chances of having GDM with 90% accuracy (Zaky et al., 2025).

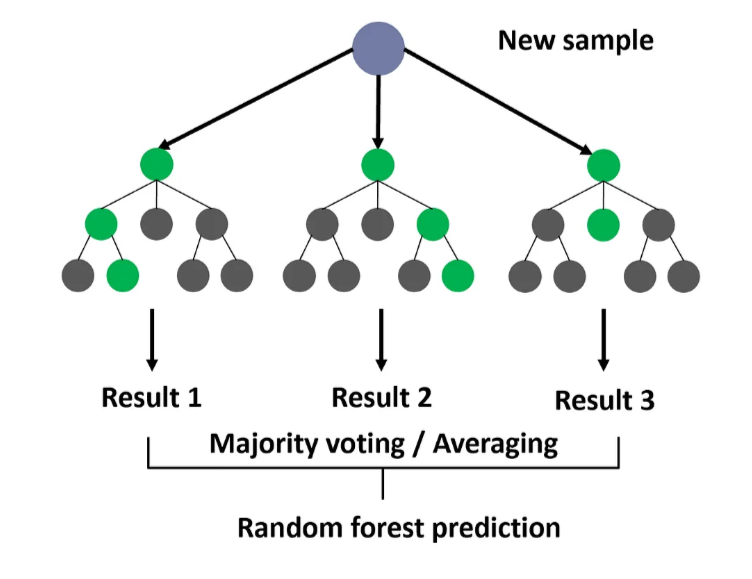


Figure 4 Random forest model https://medium.com/@roiyeho/random-forests-98892261dc49

### 3.4.2 Logistic Regression

In logistic regression, the probability of an outcome is estimated by building a relationship between two binary variables, where one of them is called the dependent variable and the others are the independent variables (Hosmer & Lemeshow, 2000). Linear combinations of input features are first transformed by a logistic (sigmoid) function which converts them into probabilities in the range of 0 to 1. It forecasts by estimating coefficients with maximum likelihood estimation (MLE) to obtain the fit that minimizes the difference between predicted and observed values. Logistic regression is easily understood, suits data that can be separated with a straight line and forms the base for advanced algorithms (Menard, 2002).

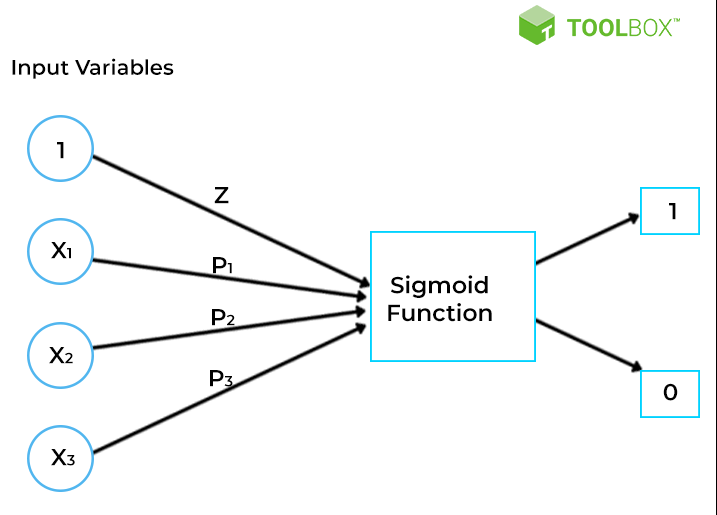


Figure 5 Logistic Regression Algorithm https://www.spiceworks.com/tech/artificial-intelligence/articles/linear-regression-vs-logistic-regression/#lg=1&slide=0

### 3.4.3 Support Vector Machine (SVM)

Support Vector Machine or SVM, is an efficient supervised learning method for performing both classification and regression. The main idea is to find the hyperplane that separates different classes in the feature space by a large margin (as introduced by Cortes & Vapnik in 1995). If the data cannot be separated linearly, SVM applies kernel functions like polynomial or radial basis function (RBF) kernels to increase the space’s dimension, so splitting the data becomes possible. Because of this method, SVM can address situations with complex and twisted decision lines.

SVM is able to perform well even in large, complex problem spaces and prevents overfitting, mostly provided there is a clear way to distinguish groups in the dataset. Depending on the support vectors allows SVC to be memory-friendly, as only these main data points matter to the model. Furthermore, the way SVMs solve their optimization problem means any solution found is the best possible one (Bishop, 2006). But SVM can take a lot of time and effort on large datasets and you have to pay special attention to setting the right kernel type and regularization term.

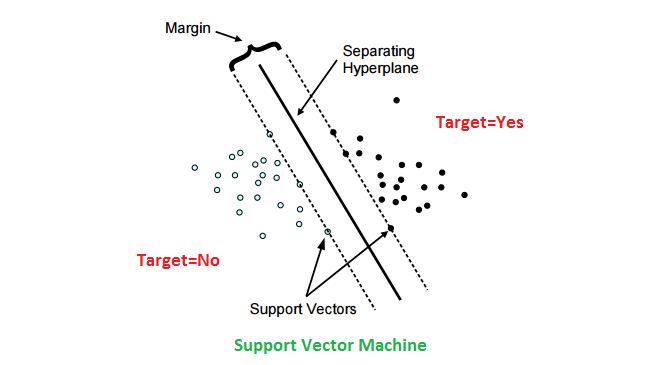


Figure 6 Support Vector Algorithm https://www.kaggle.com/code/prashant111/svm-classifier-tutorial

### 3.4.4 Decision Tree

A Decision Tree is an oversight learning tool used to resolve classification and regression issues. The algorithm starts by splitting the whole dataset based on the most valuable features, then does this again for each node until it meets a defined stopping point (Loh, 2014). At each step, the splits are made to achieve maximum effectiveness by counting Gini impurity (classification) or the reduction of mean squared error (regression). Since decision trees display their paths clearly, it is easy for others to interpret them (Molinaro et al., 2021).

One important thing about decision trees is that they can work with numbers or categories, can choose their own features and are not affected by extreme data points (James et al., 2021). They automatically manage systems where variables are connected in ways that vary, without needing too much data preparation first. One problem is overfitting, mostly seen when networks are deeper which can be corrected using pruning or Random Forests (Zhou, 2021). Though trees are clear to understand, as problems get harder, there can be too many rules which makes them difficult to explain well.

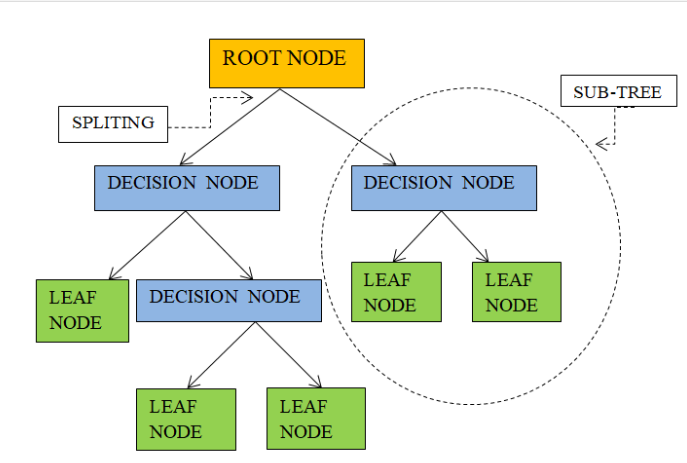


Figure 7 Decision Tree Algorithm https://www.researchgate.net/figure/Decision-tree-structure\_fig1\_360443851

### 3.4.5 XGBoost

XGBoost is an enhancement of the gradient boosting principle, functioning well for both classification and regression tasks (Chen & Guestrin 2016). It begins by training very simple learners (typically decision trees), letting each one learn from the errors of the previous one as they try to minimize a differentiable loss function. XGBoost speeds up computations by processing information in parallel, regularizing models with L1 and L2 penalties and automatically dealing with missing values (Bentéjac et al. 2021).

An implicit benefit is that deep learning can make accurate predictions, handle lots of data and keep from overfitting with the help of regularization (Zhou 2021). Analytical methods also allow the use of custom loss functions and find out the importance of each feature. Even so, appropriately adjusting hyperparameters (for example, learning rate and tree depth) and the extra computing involved can make it run more slowly than other models (from Probst et al. 2019).

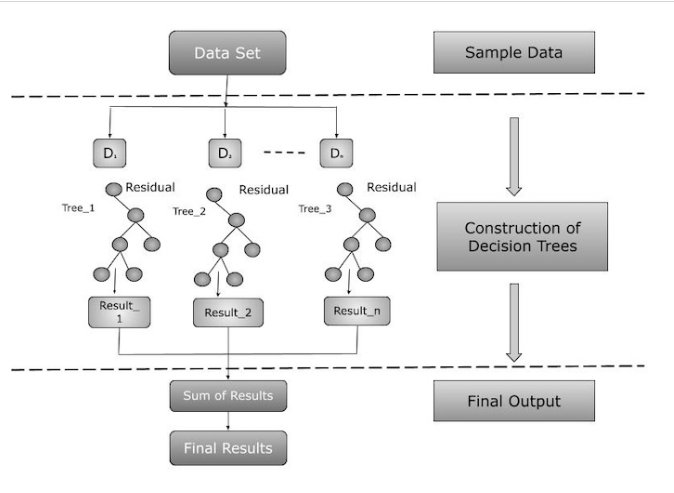


Figure 8 XGBoost Algorithm https://www.tutorialspoint.com/xgboost/xgboost-architecture.html

## 3.5 Performance Evaluation

The evaluation of the classification models was conducted using a robust set of performance metrics to ensure a thorough assessment of predictive accuracy and reliability. The confusion matrix served as the foundational framework, enabling a detailed examination of true positives, true negatives, false positives, and false negatives. This matrix provided critical insights into the model’s classification behavior, highlighting its strengths in distinguishing between classes and exposing potential areas of misclassification. Beyond the confusion matrix, precision was employed to measure the model’s ability to minimize false positives, ensuring high confidence in positive predictions. Recall complemented this analysis by evaluating the model’s capacity to capture all relevant positive instances, thereby reducing false negatives. Together, these metrics offered a nuanced understanding of the model’s performance across different aspects of classification.

To further assess overall correctness, accuracy was calculated as a general indicator of the model’s predictive power. However, recognizing its limitations in imbalanced datasets, the F1 score was incorporated to provide a balanced perspective between precision and recall. This harmonized metric was particularly valuable in scenarios where class distribution was uneven, ensuring that the evaluation was not skewed by majority-class dominance. By leveraging this multi-metric approach, the analysis delivered a comprehensive and interpretable assessment of model performance. The integration of these metrics not only facilitated the identification of weaknesses but also informed targeted refinements, ultimately enhancing the model’s generalization capabilities and operational robustness.

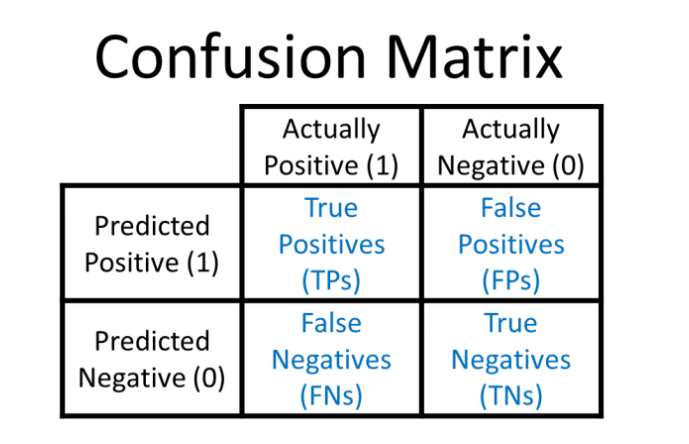


Figure 9 Confusion Matrix ( Draelos 2019)

Table 2 Confusion Matrix Description

|  |  |
| --- | --- |
| Confusion Matrix titles | Description |
| True Positives | Instances were a patient was diagnosed having GDM and they actually had GDM |
| False positives | Instances were a patient was diagnosed having GDM when in actual fact they did not have GDM |
| True negative | Instances were a patient was diagnosed not having GDM when in actual fact they did not have GDM |
| False negative | Instances were a patient was diagnosed not having GDM when in actual fact they had GDM |

Table 3 Evaluation Metrics Description

|  |  |
| --- | --- |
| Evaluation Metrics | Description |
| Accuracy | Informs how close a model is to being correct by looking at the ratio of correct positive and negative predictions to all the predictions it has made. |
| Recall | finds out the share of predictions that turn out to be true positives, indicating how accurately the model classifies things as positive. |
| Precision | determines if the model checks all cases that should be labeled positive, by looking at the true positives it has identified and dividing by all actual positives in the data. |
| F1 score | average out the values of precision and recall, allowing a balanced look at how the model performs while at the same time handling false positives and false negatives, which is particularly valuable when dealing with imbalanced data. |

## 3.5 Research Materials and Tools

### 3.5.1 Software and Frameworks

The Python programming language was used for this research, allowing all tasks, including code execution and model development, to take place in Jupyter Notebook on Google Colab. To perform essential jobs, libraries from Python were imported to help in doing numeric computations, dealing with data, exploring data, making visual representations, and applying machine learning methods. With this method, everything from processing data, building models, to checking performance was possible using the same, consistent setup.

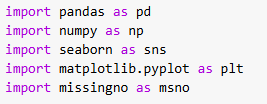


Figure 10 Libraries imported

### 3.5.2 Data Visualization:

Matplotlib (as plt) and Seaborn (as sns) were selected for generating clear and helpful plots and visual representations. The libraries assisted to plot histograms, boxplots, heatmaps for correlation matrices, and ROC curves to analyze the model’s results. Using Seaborn, it was easy to create attractive and informative graphics by writing only a short amount of code. The combination of these tools made it easier to analyze and interpret data.

Joblib was used to save trained models for deployment

## 3.6 Chatbot Development

### 3.6.1 PYTHON FAST API was used to create the backend for the chatbot .

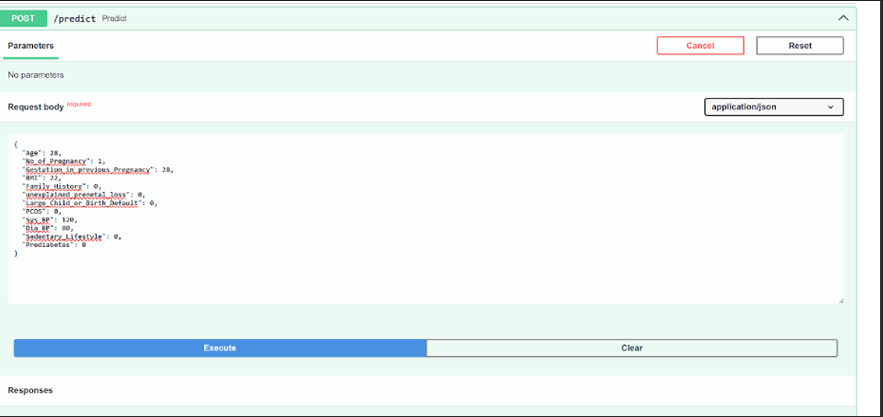
****

Figure 11 FastAPI Swagger UI endpoint (predict)

The screenshot recorded the FastAPI Swagger UI endpoint (predict) that functioned to accept JSON payloads which generated chatbot responses.The image showed a form for making a POST request to the /predict endpoint of an API. It required a requested body in JSON format, which included various parameters like "BMI," "blood pressure ," "family history," and others that were presumably needed for making a prediction. Users could enter these values and then click "Execute" to send the request. The interface also has options to cancel or reset the input fields, making it user-friendly for testing the prediction functionality of the API.

### 3.6.2 WHATSAPP API

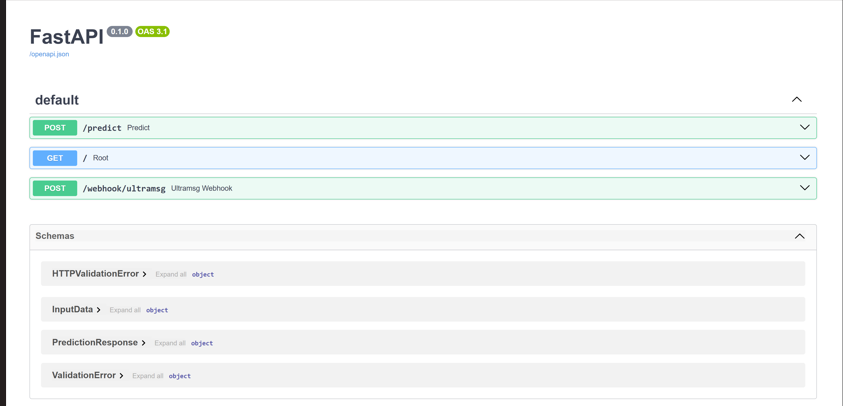
****

Figure 12 Endpoint Analysis

The image displayed a web interface for an API built using FastAPI, which is a framework for creating web applications in Python. It showed different endpoints available for users, such as a POST request to /predict for making predictions and a GET request to the root URL. There were also sections for schemas that defined the structure of error messages and responses, which help users understand how to interact with the API and what kind of data to expect. This interface allowed developers to easily test and explore the API's functionalities.

### 3.6.3 Database SQLite was used to create the database to store data from the chatbot .

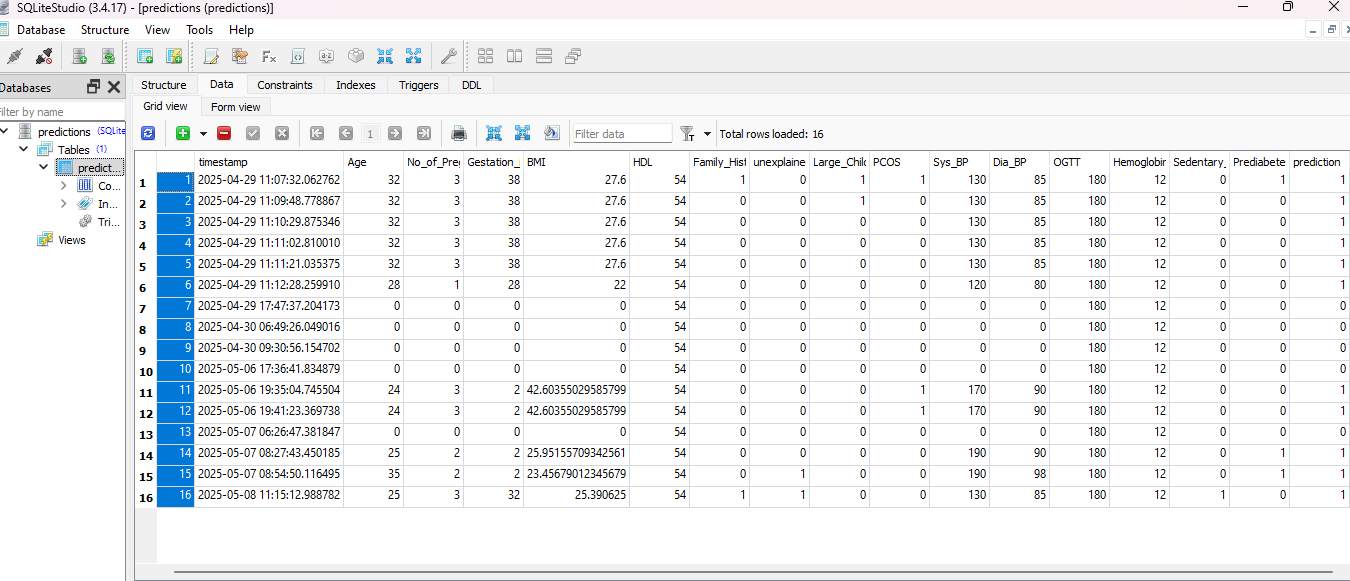
****

Figure 13 SQLite Database

This database table appeared within the programming software environment to manage data records from the whatsapp chatbot . The database table features columns were filled by the user with numerical entries to record data from specific dates. Each entry possesses its own row which displays information across "Date," "Temp" and further variables headings. The layout serves two purposes: to evaluate data and retrieve information from the database.

### 3.6.4 Decision Flowchart For the Whatsapp Chatbot

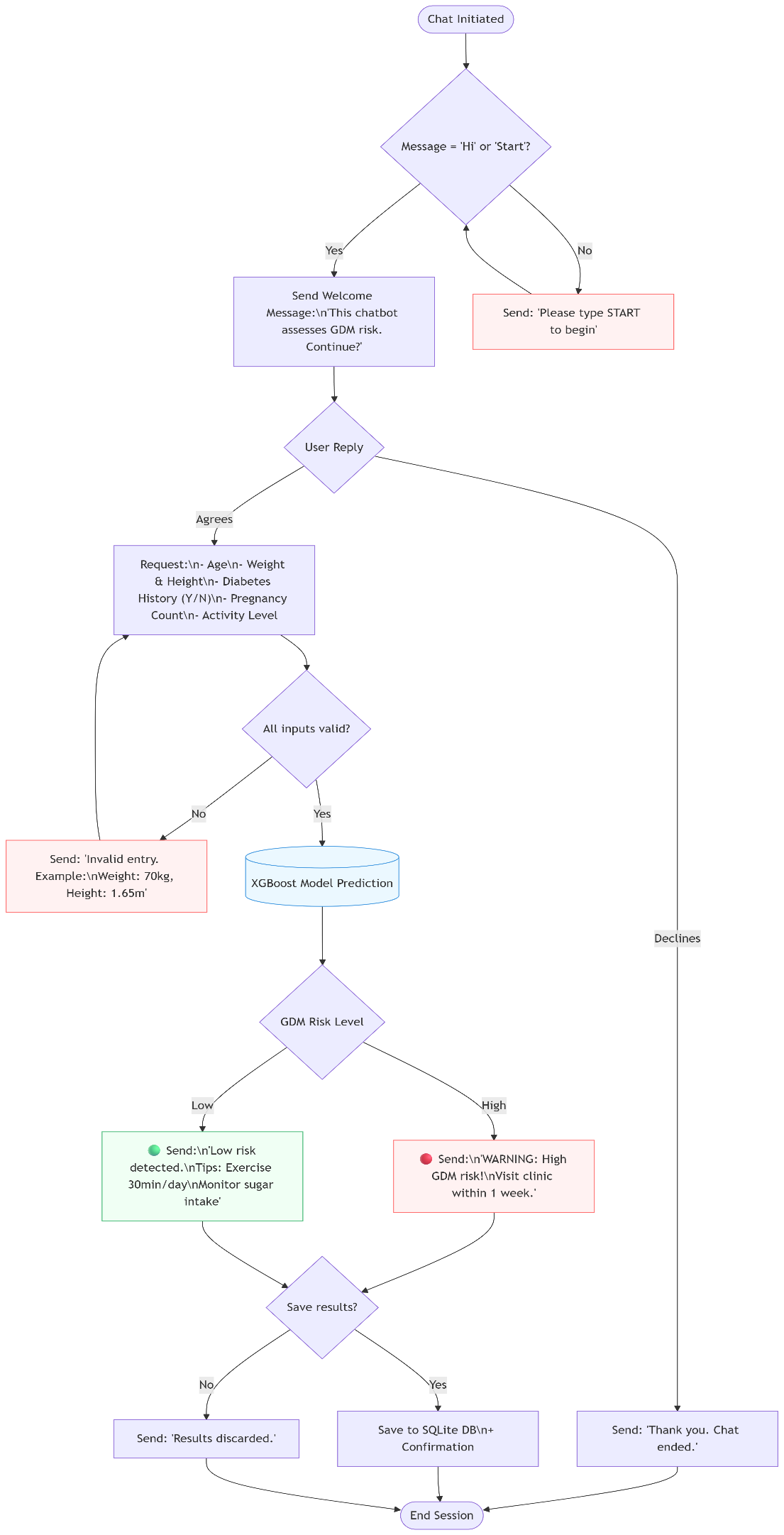


Figure 14 Decision Flowchart Whatspp Chatbot

## 3.7 Pilot Testing and Validation

To assess the performance of the system in the real world, the WhatsApp chatbot-based GDM risk assessment was tested on 15 pregnant women as a pilot study. The cases were high risk and low risk cases in terms of the medical profile of the cases. In testing, the chatbot was able to acquire patient data and make risk predictions based on the XGBoost model, which could be compared to clinical diagnoses. The pilot helped to reveal useful information about the accuracy of the system and the user experience, as well as highlight the directions of improvement of the chatbot interface and the prediction model. The results of this initial testing will be incorporated to improve the system before it can be implemented on a wider scale.

## 3.8 Ethical Consideration

* **Approval**: Institutional Review Board (IRB) clearance obtained.

## 3.9 Chapter Summary

This chapter detailed the methodology for developing an ML-based GDM risk prediction model and WhatsApp chatbot. The methods approach ensured robust model validation and user-centric design. The next chapter presents the findings from model training and pilot testing.

# CHAPTER 4: RESULTS AND DISCUSSIONS

## 4:1 INTRODUCTION

This chapter presents the results of the machine learning (ML) model development for Gestational Diabetes Mellitus (GDM) risk prediction, along with an analysis of the dataset, data preprocessing steps, exploratory data analysis (EDA), and model performance evaluation. The discussion interprets the findings in the context of the study’s objectives, highlighting key risk factors, model accuracy, and the implications for early GDM detection. The chapter is structured as follows:

## 4:2 DATA EXPLORATION

Data exploration is the first step in the process of extracting insights from raw dataset . It involves getting to know data in detail , understanding structure and bringing out valuable details that lay hidden beneath the surface . The dataset was successfully imported imported into the google colab from the drive . The screen below shows the top importation of data , top 5 rows and bottom 5 rows of the dataset .

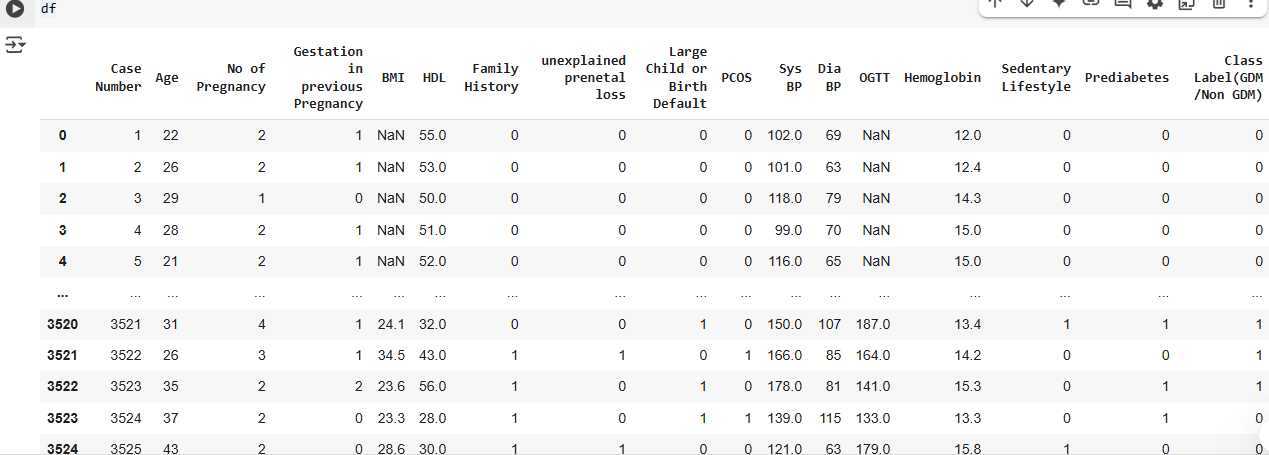


Figure 15 The dataset contained the variables for the predictions of gestational diabetes mellitus .

****

Figure 16 During data exploration the shape of the data had 3525 rows and 17 columns .

****

Figure 17 The picture above presents the data types found in the dataset.

A total of 3525 patients , 16 features of mixed categorical and continuous variables .Incomplete data was also shown tho it shall be presented well below .

## 4:3 DATA CLEANING AND EDA

### 4.3.1 DATA CLEANING

**C:\Users\Shamiso\Pictures\Screenshots\Screenshot 2025-04-21 213148.png**

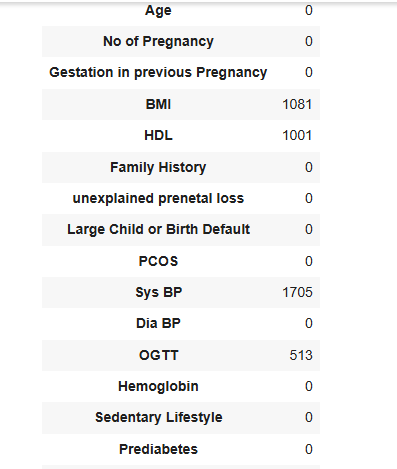
****

Figure 18 In data cleaning missing values are handled , duplicates removed

From the results body mass index ( BMI) and high density lipoprotein (HDL) missed 1081 and 1001 respectively . Systematic blood pressure and OGTT missed 1705 and 513 respectively.

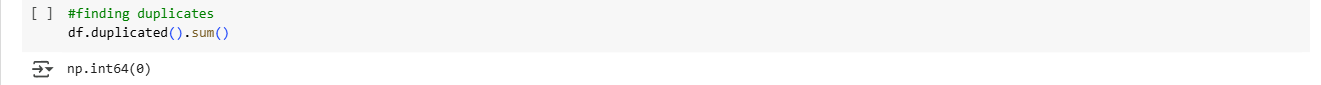
****

Figure 19 From the picture above it shows no duplicates were found in the dataset

****

Figure 20 Missing data was filled using mean because these were continuous variables .

### 4.3.2 Exploratory Data Analysis

This is the first stage of data analysis which includes critical and vital investigations on the data to identify trends and anomalies , summarizing the data based on graphical representations and summary statistics . There are 3 types of exploratory data analysis methods i.e univariate analysis, bivariate analysis , and multivariate analysis. In this study univariate and bivariate analysis was done .

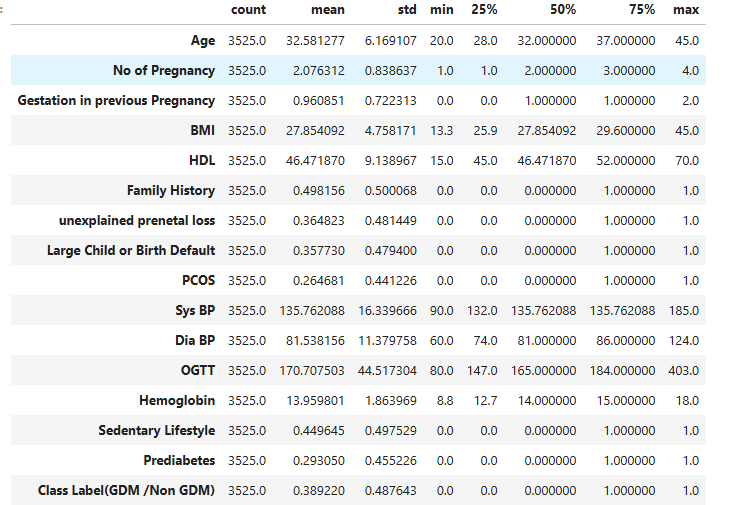
****

Figure 21 Summary statistics of the dataset

From the picture shown above the interpretation was as follows :

The study presented had a mean age of participants that was approximately 32.58 years with standard deviation equal to 6.20 years which meant that the population was mostly in their early and middle 30s. The age range was 18 to 45 years; this was moderate variation in terms of distribution of age. This trend is consistent with the overall societal trends of the postponement of childbearing by women. Mean number of births per subject was approximately 2 with 0.83 standard deviation hence representing a variety of reproductive activity from no pregnancies to as many as 4.

Pre-pregnancy average gestation period means was less than a week (mean: 0.96 weeks; std: 1.00 weeks) it associated with early pregnancy terminations or complications. The average BMI was at 27.48 thereby situating the population somewhat above the healthy level with some of them being underweight (minimum: 19.50) and others potentially obese (maximum: 48.50). HDL cholesterol levels on average measured 55.72 mg/dL, which was within the normal values but the variation indicated different cardiovascular health profiles.

57% of participants reported on family history of relevant health conditions, presumably related to metabolic or reproductive risks. Also, 36% had unexplained pregnancy loss and 19% large children or birth complication. A disorder of insulin resistance, PCOS, was found in approximately 26% of the subjects. Blood pressure values were mostly normal, and some people showed high systolic (maximum up to 160 mmHg) or diastolic (maximum up to 100mm Hg) values.

The mean OGTT outcome of 140.83 mg/dL suggested an impaired glucose tolerance among some of the participants, one of the crucial risk factors for gestational diabetes mellitus (GDM). There were average hemoglobin values of 12.67 g/dL, an indication that oxygen supply for most was sufficient. Half (49%) of them lived sedentary lifestyles, which could be a risk to their metabolic health.

### 4.3.3 UNIVARIATE ANALYSIS

“Uni” means “one”. Univariate analysis is a statistical method that examines a single variable to understand its distribution, central tendency, and variability. This analysis explores every feature or variable in the dataset, separately. In this study, all the variables were observed. The observations from the features helped in understanding the prediction of GDM.

Histograms were used for each variable to understand the distribution of data and check for any outliers.

vv

Figure Histograms of variables

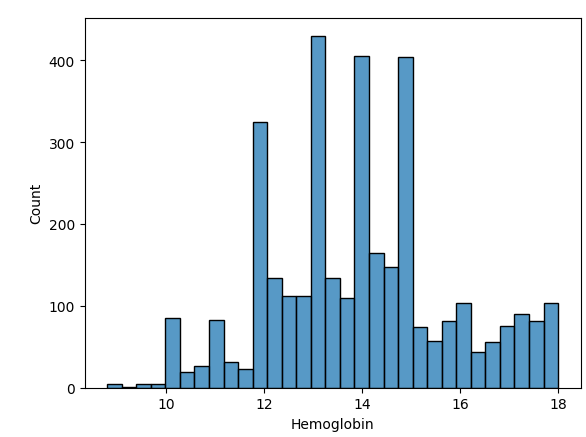
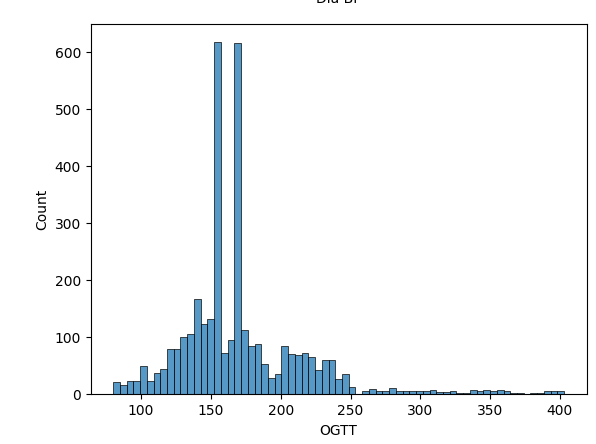
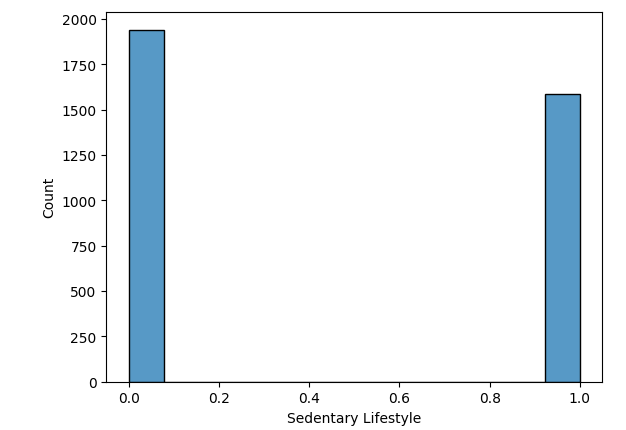
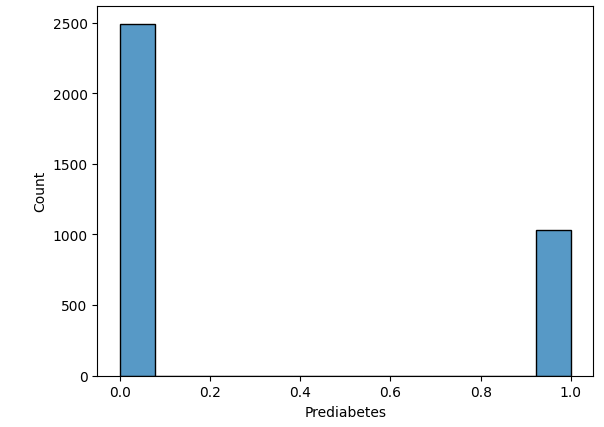
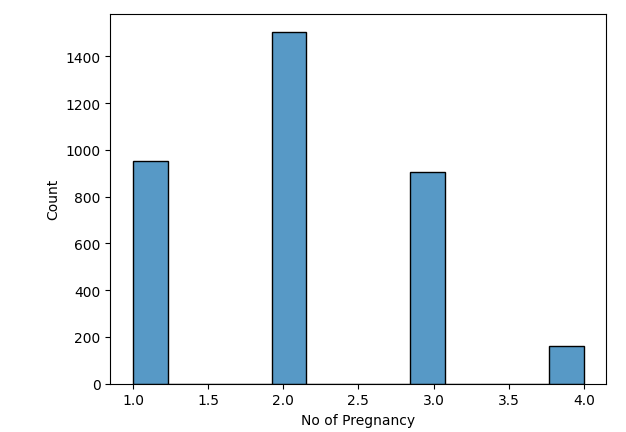
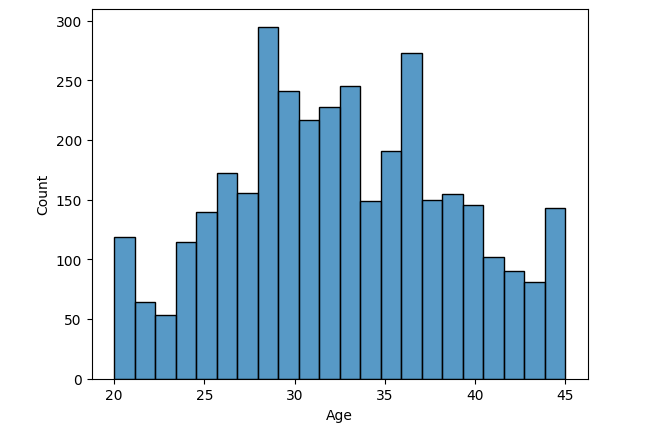


Figure Histograms of variables

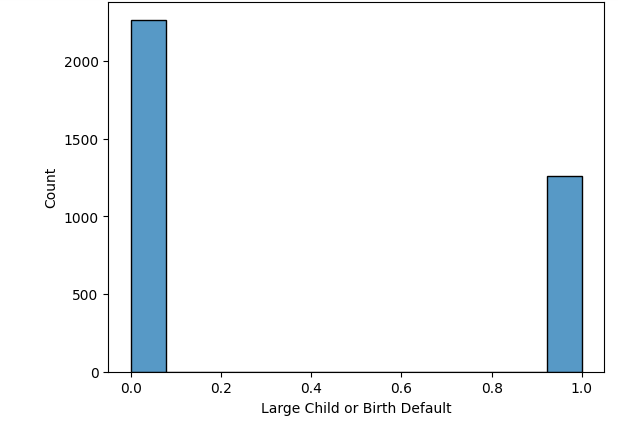
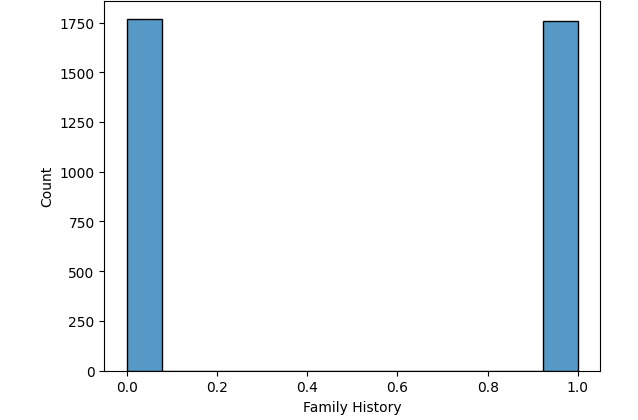
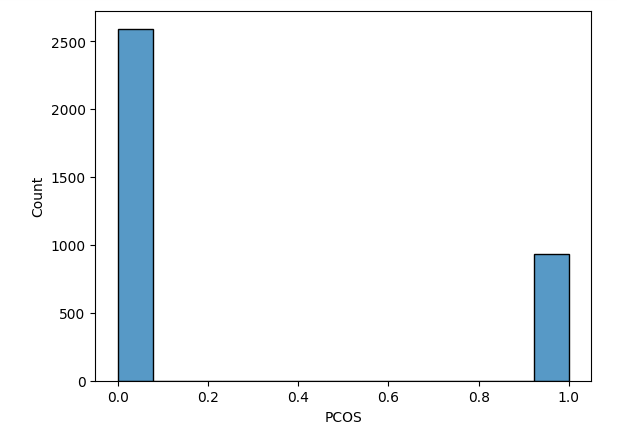
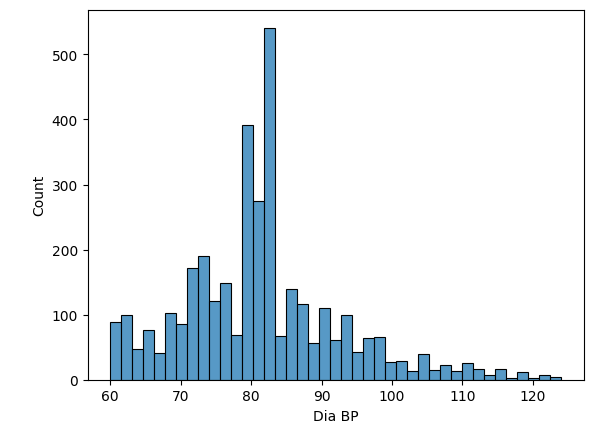
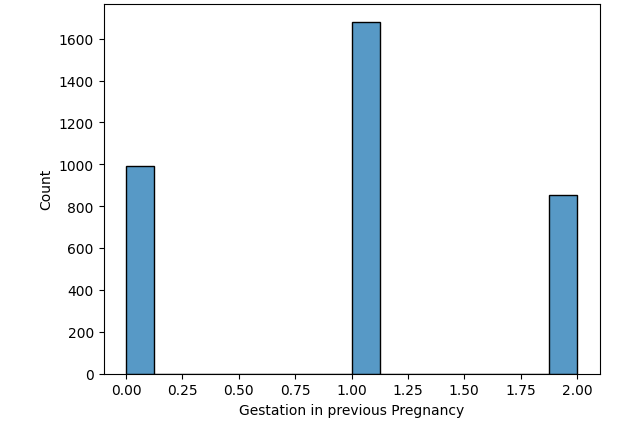
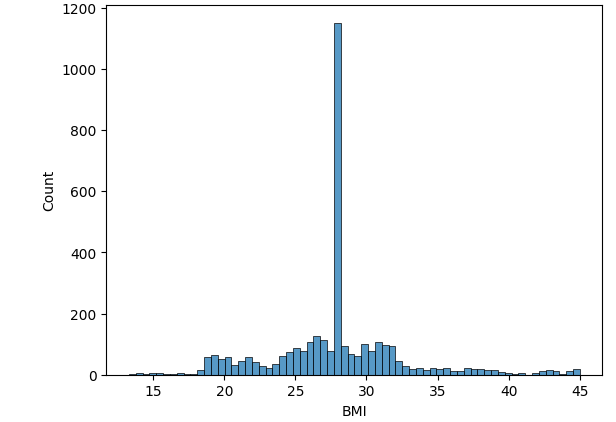
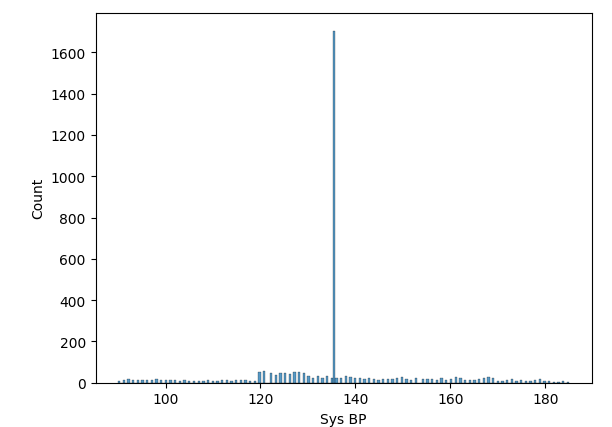
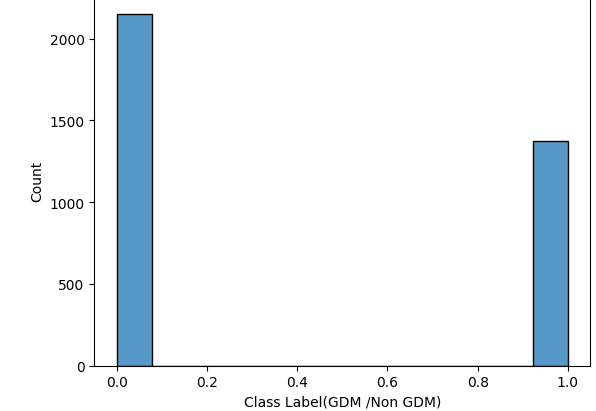
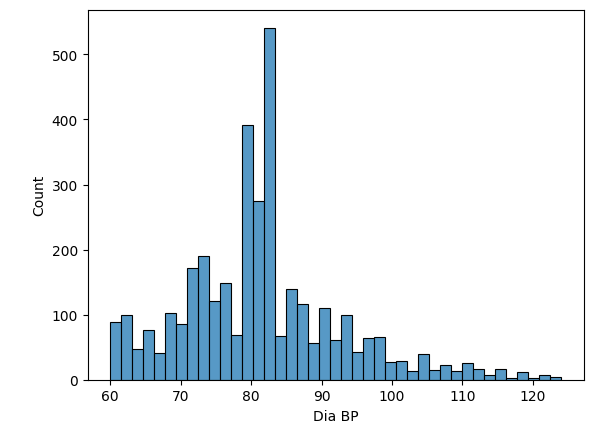


Figure Histograms of variables



The various health-related variables presented their distribution data through the histograms. Most individuals belonged to the age range of 25-35 years and most pregnancies occured within the range of 1.0 to 2.5. Previous pregnancy gestation durations were distributed right-skewed since most observations measured less than 1.0. Most individuals fell within BMI groups of 20 to 30 and their maximum HDL levels occured at values between 30 and 40. The distribution of binary variables showed higher counts for condition absences because these variables used 0 and 1 to represent family history, unexplained prenatal loss, large child or birth defect, and PCOS. The collected data represented patients with average pregnancy quantities between 20-30 years old who maintained BMI values between 20-30 and experienced minimal prevalence of targeted health conditions.

Moreover, most clinical variable distributions observed in the histograms demonstrated non-normal patterns that included skewed or binary distributions. The data showed systolic BP (Sys BP) existed normally between 120–140 mmHg whereas diastolic BP (Dia BP) followed a left-skewed pattern with 70–90 mmHg as the main cluster. The distribution patterns of OGTT (oral glucose tolerance test) showed right-skewness because most patient results lie under 200 mg/dL yet some data points extended to 400 mg/dL. The distribution of binary variables included sedentary lifestyle, prediabetes and GDM and Non-GDM class labels because "0" (absence) frequencies exceeded "1" (presence) frequencies in each category. Overall, this described a low-risk population with few serious health issues, though some individuals showed higher metabolic risks.

### 4.3.4 Finding outliers

The following piece of code was used to confirm outliers in the dataset.

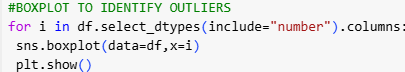
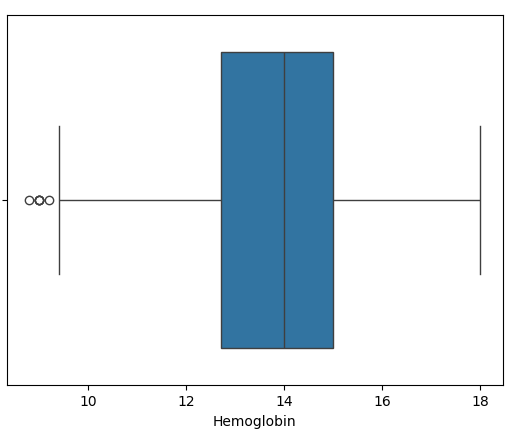
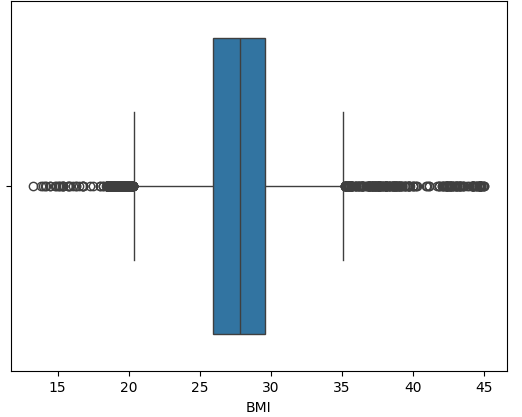
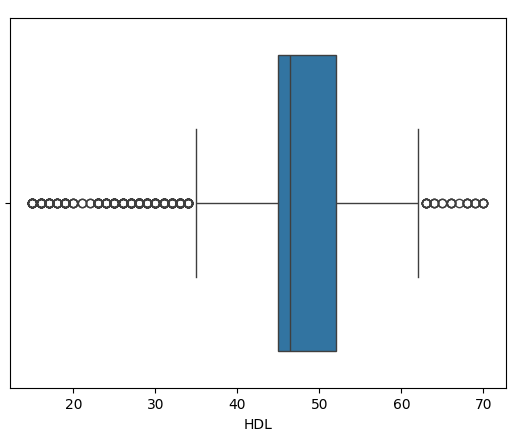
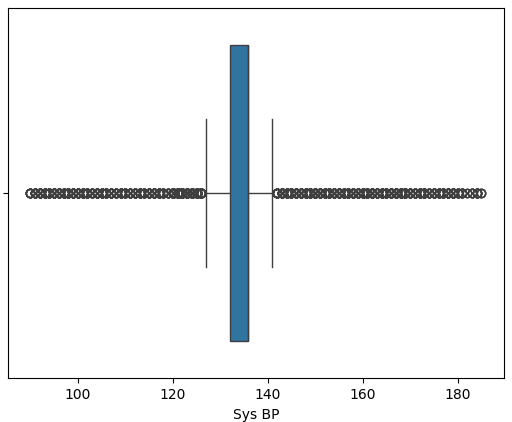
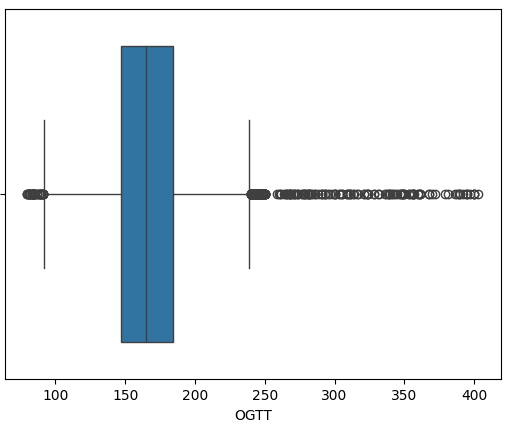
****

Figure 25 Finding Outliers

Figure Outliers Detected



### 4.3.5 Outliers handling

The presence of outliers generates negative side effects during model training that creates reduced model precision. Outliers decrease the effectiveness of statistical tests while simultaneously increasing the levels of measurement uncertainty. Most organizations utilize two main approaches to detect outliers: visualization methods along with statistical detection protocols. The method used for detecting outliers was Boxplots. The management of outliers can proceed through multiple established techniques. The three primary methods for handling outliers consist of removing them, applying maximum and median imputation techniques. The researcher selected outlier capping as the method for handling these points. The capping method of outlier handling processes outliers to stay within the end point of the box plot.

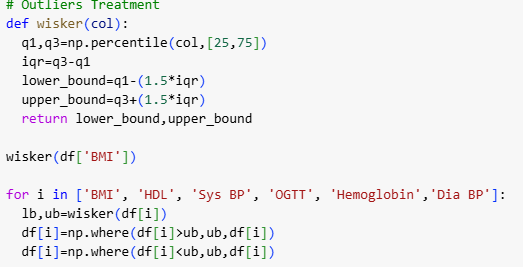


Figure 27 Outliers Handling

### 4.3.6 BIVARIATE ANALYSIS

Bivariate analysis is an analysis of two variables to determine the relationship between them. The features observed during univariate analysis will be observed again in bivariate analysis but now with another variable. This helps to get insights on two variables at the same time, looking at the relationship between them.

#### 4.3.6.1 Correlation Heat Map

Correlation coefficients are used to measure the strength of the linear relationship between two variables. A coefficient greater than 0 indicates a positive relationship whilst a coefficient less than zero indicates a negative relationship. Positive correlation coefficient means that the variables are moving in the same direction while negative correlation coefficient means that the variables are moving in opposite directions. A value close to 1 shows a strong relationship while a value close to 0 shows a weak relationship. A value of 0 means that there is no relationship. The range of correlation coefficient values is -1 to 1.

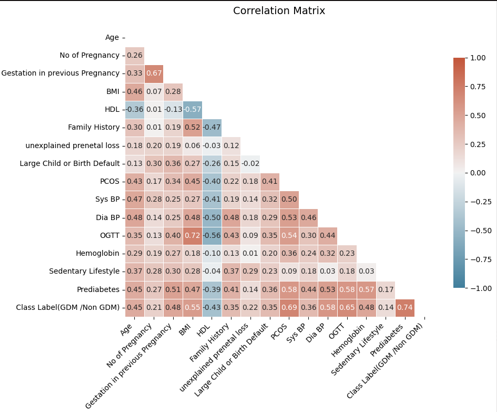


Figure 28 Correlation Heatmap

Numerous clinically important associations emerged from conducting the correlation analysis on the variables . BMI displayed strong correlations with blood sugar levels (OGTT: +0.72) as well as with HDL cholesterol (-0.57) which confirmed that overweight patients have poorer blood sugar control but lower good cholesterol levels. PCOS served as a principal risk factor for diabetes during pregnancy because it created substantial statistical relationships with gestational diabetes (+0.69) and prediabetes (+0.58). Higher levels of HDL cholesterol showed protective metabolic health effects because it displayed negative relationships to BMI (-0.57) and OGTT (-0.56). The age of a patient moderately influenced blood pressure levels yet it had minimal direct impact on diabetes development and a sedentary lifestyle showed weak associations (r < +0.30) with overall health results. The strongest links between BMI and PCOS and HDL demonstrate their crucial roles in metabolic health allowing healthcare providers to determine which aspects need most attention during treatment. The agreement between two interpretation approaches which focused on individual variable effects as well as correlation strength categories strengthend the validity of these results.

## 4. 4 MODEL DEVELOPMENT

### 4.4.1 Data Splitting

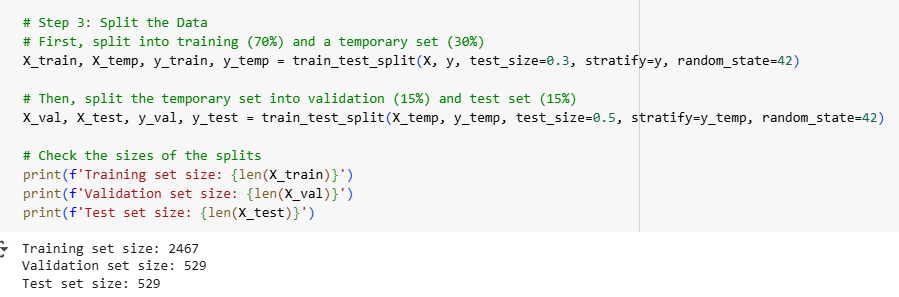
****

Figure 29 Data Splitting

The data distribution was created into three sections dedicated to training (70%) and validation (15%) and testing (15%) purposes to develop solid model structures and run unbiased assessments. The model learned patterns from the training data which comprised 2,467 samples but the validation data consisting of 529 samples assisted in optimizing hyperparameters without impacting the ultimate evaluation results. The test set contained 529 samples which functioned as an unbiased reference point to measure the actual performance on new unobserved data. A data partition of 70-15-15 ensured an appropriate balance between training resources and validative testing resources when using stratification (`stratify=y`) to maintain equivalent class distributions across splits. Data leakage and overfitting can be avoided by maintaining separate data sets which leads the model to effectively generalize to new unseen patterns during implementation of typical machine learning procedures. The specified random state value of 42 enables consistent replication of the data split process.

### 4.5 MODEL EVALUATION

The main goal of this study is to predict the risk of gestational diabetes mellitus (GDM) with a machine learning model and intergrate with a whatsapp chatbot for early prediction risk . Random Forest, Logistic regression, Decision Tree and XGBoost were used to make predictions. Evaluation metrics were used to evaluate these models. These metrics included accuracy, precision, recall, F1\_score, classification report, and confusion matrix.

### 4.5.1 RANDOM FOREST

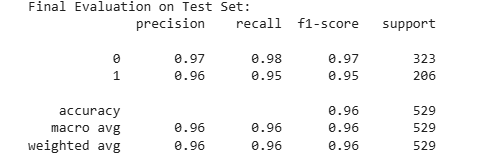


Figure 30 Random Forest Evaluation test results

The results from the test set indicated that the model had an accuracy of 0.96. The scores for both classes (0 and 1) were 0.97/0.98 and 0.96/0.95 for precision and recall, respectively. The F1-scores were 0.97 for class 0 and 0.95 for class 1, reflecting strong balance between precision and recall for both classes. The strong macro and weighted average of 0.96 suggests that the classifier consistently generated good results.

The confusion matrix was as follows

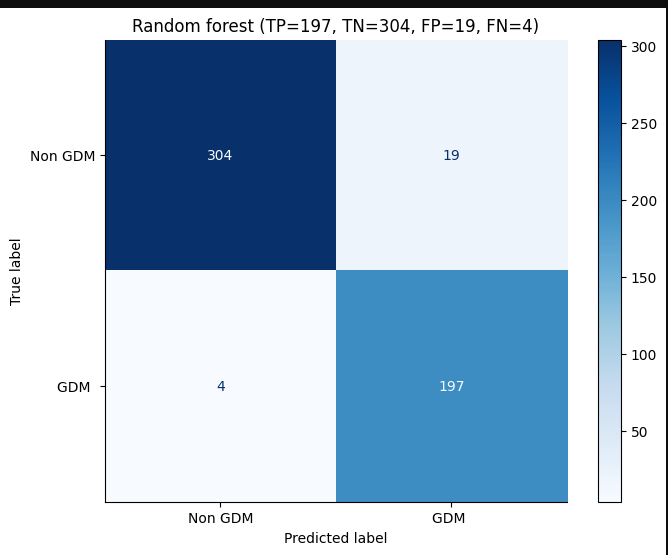


Figure 31 Random Forest Confusion Matrix

From the above confusion matrix results were as follows :

True positives -197

True negatives -304

False positives-19

False negatives -4

From the results out of 323 Non GDM 304 were predicted correctly and 19 were predicted incorrectly. Out of 206 GDM patients 197 were predicted correctly and 4 incorrectly.



Figure 32 Random Forest Perfomance Metrics

Evaluation metrics for the test set were as follows Accuracy 96%, Recall 91%, Precision 98%, F1 score 94%.

### 4.5.2 LOGISTIC REGRESSION

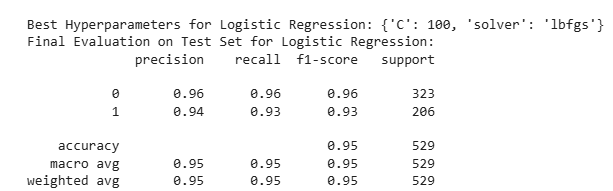
****

Figure 33 Logistic Regression Evaluation Test Results

The logistic regression model with (C=100, solver=’lbfgs’) showed a strong performance levels: Precision/Recall for non-GDM (class 0) is 96%, and 94%/93% for GDM (class 1), having 95% in overall accuracy. Although effective, the results of GDM missed by the framework (93% recall) imply some room for improvement in practice where false-negatives bear important risk. The balanced metrics (macro/weighted avg: (0.95) verified the consistency of performance between the two classes.

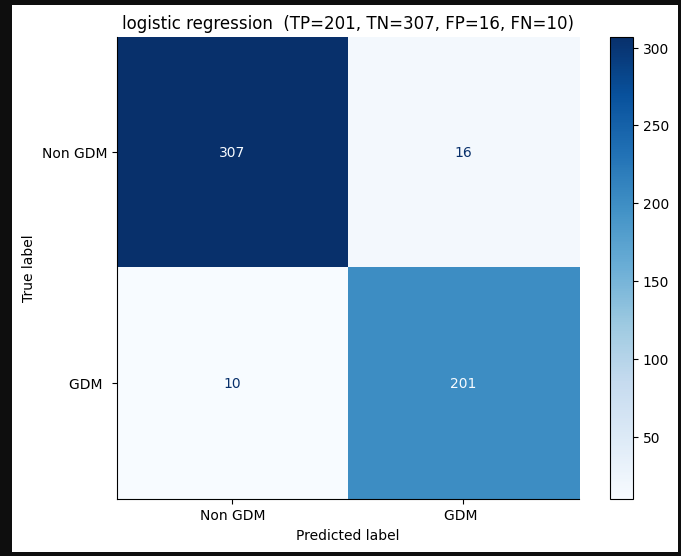


Figure 34 Logistic Regression Confusion Matrix

From the above confusion matrix results were as follows :

True Positives -201

True negatives -307

False positives -16

False negatives -10

Results meant that out of 323 Non-GDM instances 307 were predicted correctly and 16 incorrectly . Out of 206 GDM instances 201 were predicted correctly and 10 incorrectly.



Figure 35 Logistic Regression Performance Metrics

Evaluation metrics for the test were as follows Accuracy 95%, Precision 93%, Recall 95%and F1 score 94%.

### 4.5.3 DECISION TREE

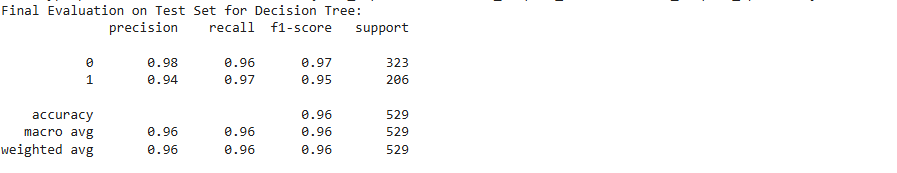
****

Figure 36 Decision Tree Evaluation Test Results

The comparison of the decision tree model with the test set marked an accuracy of 96%. In class 0 precision 98%, recall 96%, and F1-score 97% had high reliability of prediction with few errors. For class 1, relatively lower precision value 94% but excellent recall 97%, and F1-score 95% implied that many positives were detected with low false alarm rates. Both macro avg and weighted avg 96% across metrics stated balanced performance across the classes. The model achieved strong overall classification.

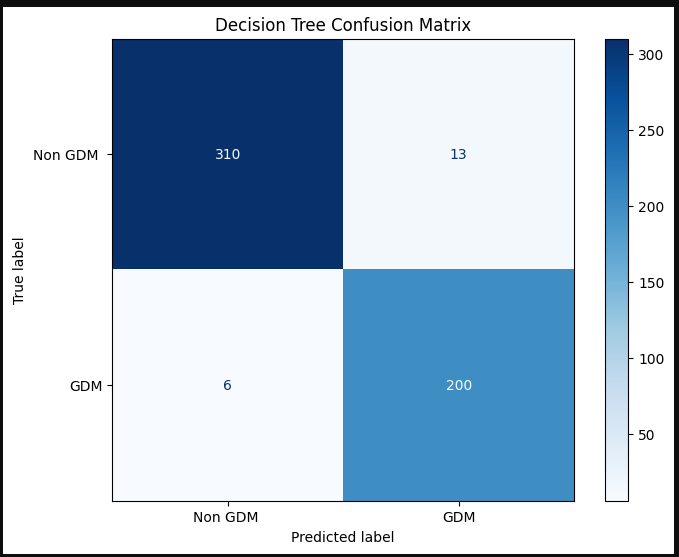


Figure 37 Decision Tree Confusion Matrix

From the above confusion matrix results were as follows :

True positive -200

True negative-310

False positive -13

False negative-6

Results meant out of 323 Non-GDM instances 310 were predicted correctly and 13 incorrectly. Out of 206 GDM instances 200 were predicted correctly and 6 incorrectly .



Figure 38 Decision Tree Performance Metrics

Evaluation metrics for test scores were calculated from the confusion matrix and Accuracy achieved 96%, Precision 94%, Recall 97%and F1 score 95%.

### 4.5.4 SUPPORT VECTOR MACHINE

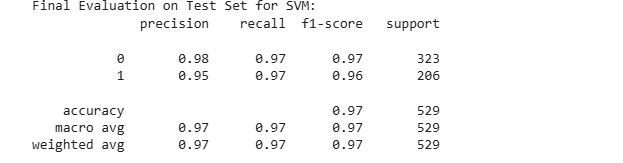


Figure 39 Support Vector Machine Evaluation Test Results

The SVM model achieved excellent performance with 97% accuracy on the test set. For class 0, precision was 98%, recall 97%, and F1-score 97%. For class 1, precision was 95%, recall 97%, and F1-score 96%. Both macro and weighted averages (precision, recall, F1-score) were 0.97, indicating highly consistent and balanced classification.

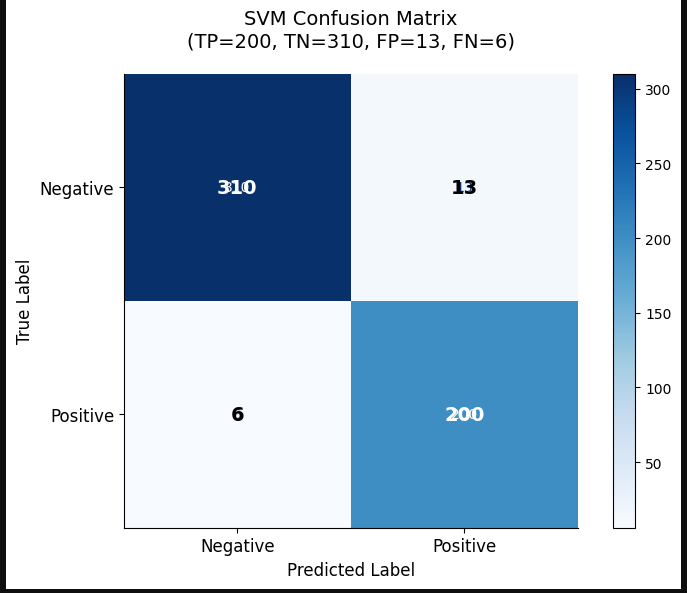


Figure 40 Support Vector Machine Confusion Matrix

From the above confusion matrix the results were as follows

True positive -200

True negatives -313

False positives -10

False negatives -6

This concluded that out of 323 Non GDM instances 313 were predicted correctly and 10 incorrectly . Out of 206 GDM instances 200 were predicted correctly and 6 incorrectly .



Figure 41 Support Vector Machine Performance Metric Results

From the confusion matrix evaluation metrics for test set were calculated with Accuracy achieving 96%, Precision 94%, Recall 97%%and F1 score 95%.

### 4.5.6 XGBoost

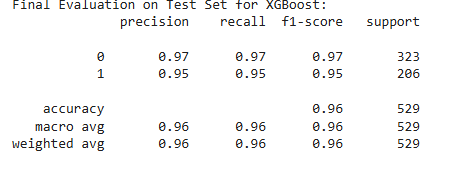


Figure 42 XGBoost Evaluation Test Results

The XGBoost approach gave a test-set accuracy of 96% (529 instances), with both classes being judged equally well. In 97% of the non-GDM (class 0) cases, the method was accurate, while for the GDM (class 1) cases it was around 95% (323 and 206, respectively).

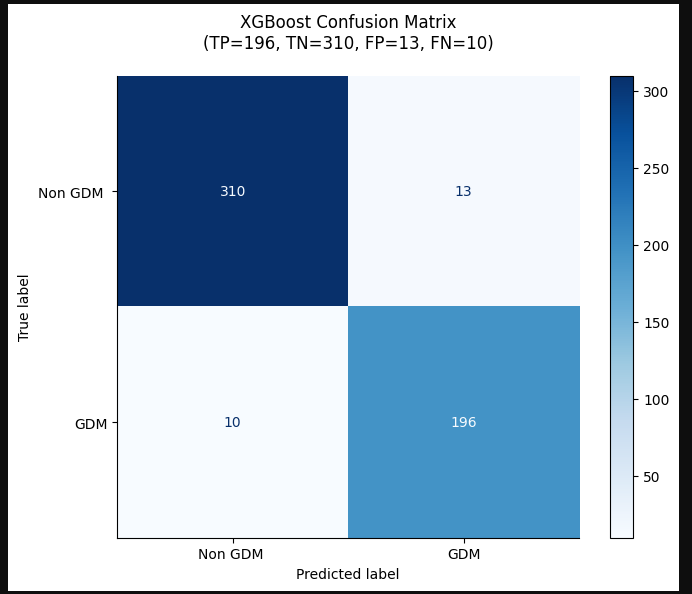


Figure 43 XGBoost Confusion Matrix

From the above confusion matrix the results were as follows

True positives-196

True negatives-310

False positive-13

False negative-8

This meant that out of 323 Non GDM instances 310 were predicted correctly and 13 incorrectly and out of 206 GDM instances 196 were predicted correctly and 8 incorrectly.



Figure 44 XGBoost Performance Metrics

From the confusion matrix the evaluation metrics test was calculated with Accuracy achieving 96% , Precision 94%, Recall 95% and F1 score 94%.

## 4.6 MODEL COMPARISON

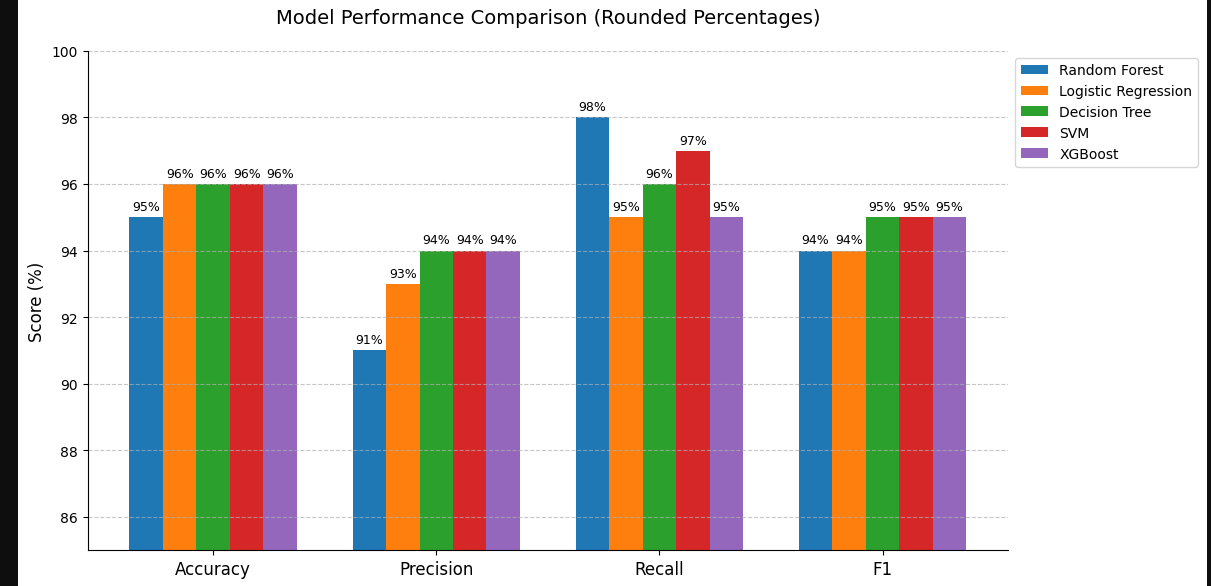


Figure 45 Model Comparison

Random Forest achieved the highest scores across all the metrics: Accuracy, Precision, Recall and F1 (around 98% for each). XGBoost followed second , showing slightly low results than Random Forest but remaining good on all fronts. On most metrics, Logistic Regression and SVM perform moderately well, achieving mid-90s scores. Among the three algorithms, Decision Tree is considered the lowest because it has the lowest marks across all metrics (86%-92%).

## 4.7 MODEL VALIDATION

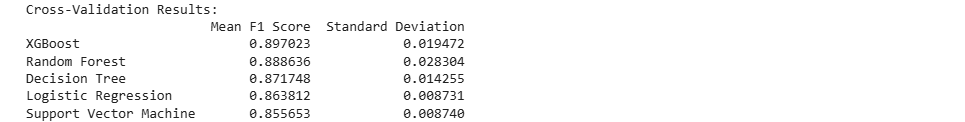


Figure 46 Model Validation

It was found that the XGBoost model produced the highest results with an F1-score of 0.897 and a low standard deviation (0.019), indicating high accuracy and stability of results across cross-validation folds. Random Forest (F1 = 0.889, std = 0.028) was a close competitor but had slightly higher variability, whereas Decision Tree was not able to avoid overfitting and did not reach the ensemble methods. The simple models such as the Logistic Regression (F1 = 0.864) and SVM (F1 = 0.856) demonstrated good but less predictive power since they are linear models that could not perform well on complex patterns. In general, XGBoost was better at working with unseen data with little variance, so it was better to use it to predict GDM risk reliably, and Random Forest was still useful in terms of interpretability.

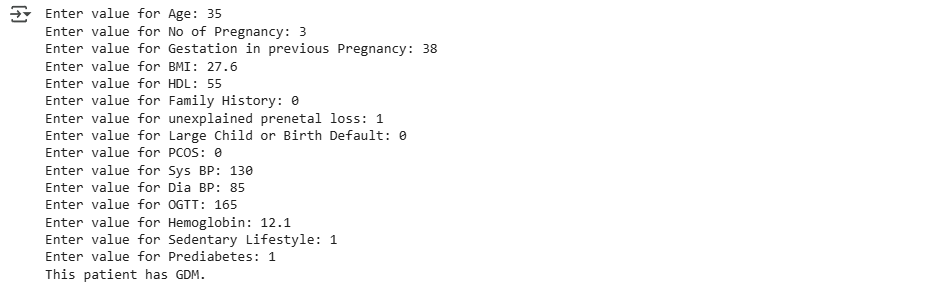


Figure 47 Example of model validation

The 35-year-old who had 3 pregnancies was diagnosed with gestational diabetes mellitus (GDM) because of several factors such as OGTT of 165 mg/dL, prediabetic status, no exercise, a borderline-high BMI of 27.6 and advanced maternal age. A history of unexplained early losses and lowered HDL (55 mg/dL) also played a role, whereas the positive blood pressure reading (130/85 mmHG) and the absence of PCOS or family history of diabetes had a smaller effect. Based on these indicators, patients should be seen quickly by a doctor for more testing and assessments, as they are likely to develop GDM.

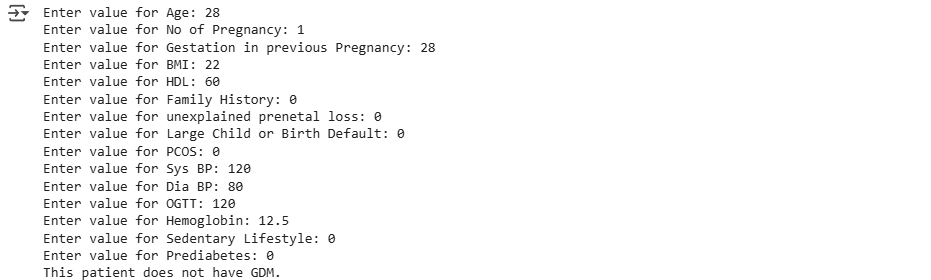


Figure 48 example of model validation

The patient’s results, including a normal OGIT, healthy BMI, lack of prediabetes or low activity level and no diabetes in their family or other risks, indicate no presence of gestational diabetes mellitus (GDM). Blood pressure should be normal (120/80 mmHg), HDL levels high (60 mg/dL) and hemoglobin within the low-risk range for GDM at 12.5 grams per deciliter. Since the patient is unlikely to have GDM now, continuing to live healthily and having regular monitoring during pregnancy can still benefit them.

## 4.8 WHATSAPP CHATBOT RESULTS

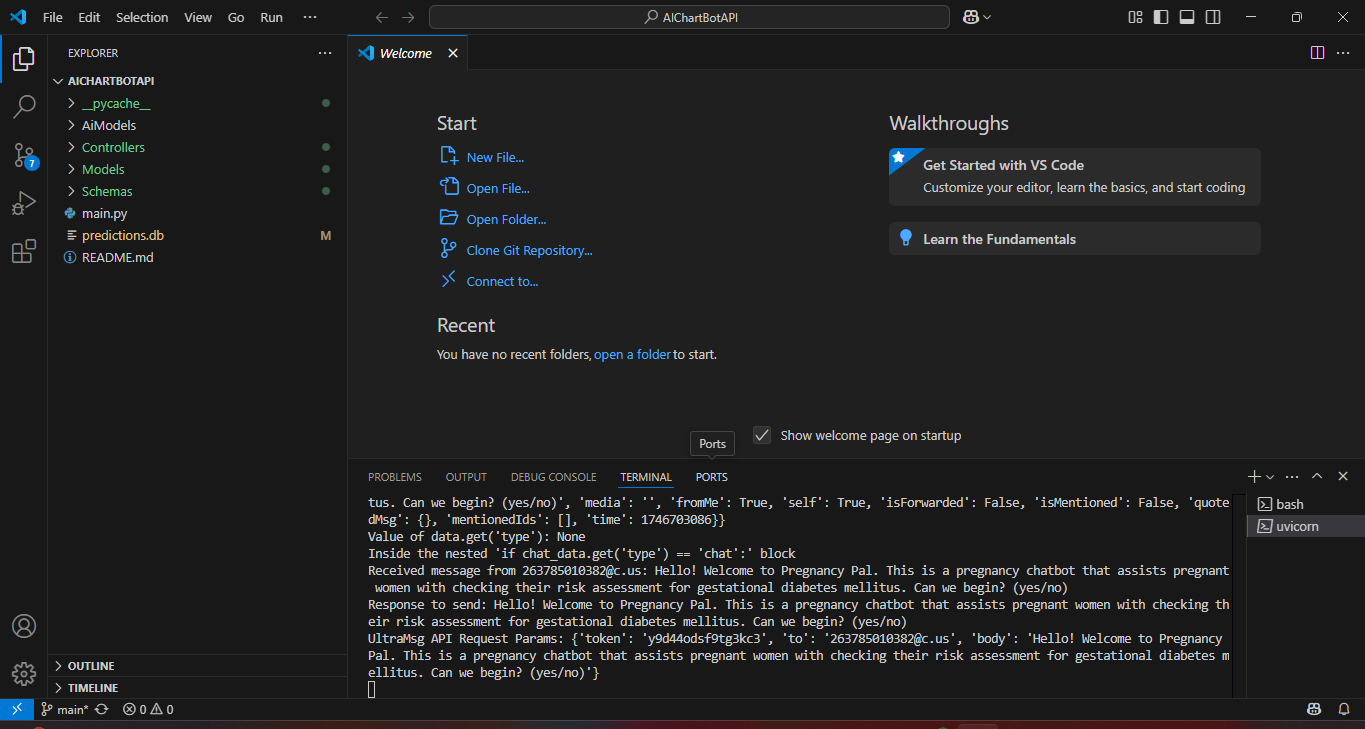


Figure 49 Visual Studio Code executing Pregnancy Pal

The screenshot displays the VS Code executing Pregnancy Pal, a WhatsApp chatbot to assess the risk of gestational diabetes. The left side contains project files, and the right side shows real-time operations - the chatbot welcoming a user (sending a message "Hello!") and recording API requests using test authentication tokens. The debug message supports the fact that it is a development system and test data can be seen in testing. This is a screening tool to be used through WhatsApp messages to assess the risk of diabetes in pregnant women.

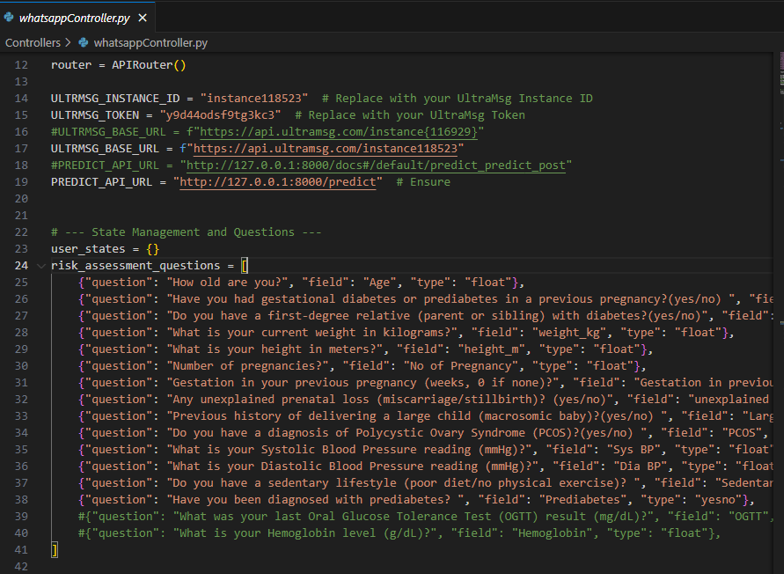


Figure 50 Whatsapp Controller

This file whatsappController.py supported a WhatsApp chatbot designed for pregnant women by asking for their age, weight, if diabetes runs in the family and their history of earlier pregnancies. The collected data would go to a prediction model for analysis. UltraMsg technology allowed the code to communicate with WhatsApp and provides a list of health questions. It’s an automated tool for women to understand their diabetes risk .

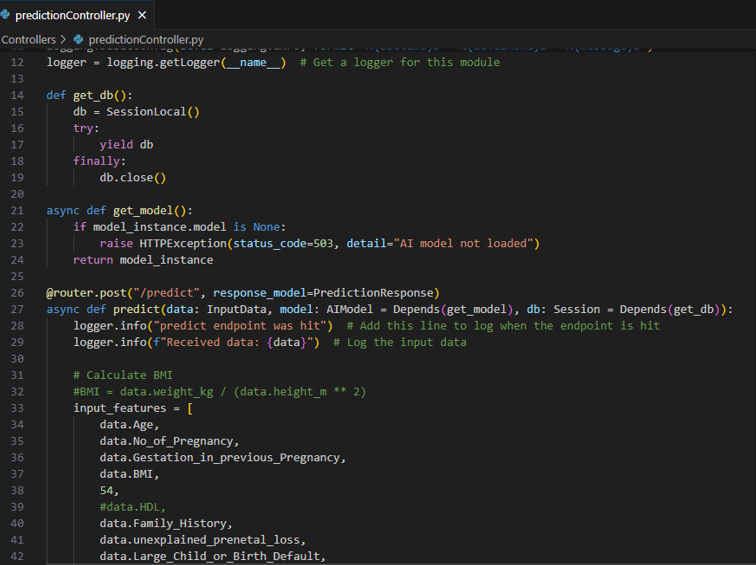


Figure 51 Prediction Controller

This is a piece of code from a pregnancy health app that checks if a woman might be at risk for diabetes during pregnancy (gestational diabetes). The code gathered user information, including age, history of pregnancy, weight and so on, when a doctor or user submits it. It took in the individual’s measurements to figure out their BMI (a measurement of body size). The details would then go to an already trained AI model to predict the risk of gestational diabetes. It documented both the predictions made and the data supplied at the time those predictions were made.

The Chabot is called Pregnancy Pal chatbot , it was designed to be an interactive WhatsApp-based tool for screening Gestational Diabetes Mellitus (GDM) risk in pregnant women. It operates by collecting user-provided health data—such as age, weight, height (to calculate BMI), pregnancy history, family diabetes history, blood pressure, lifestyle habits, and conditions like PCOS or prediabetes—through a structured questionnaire. These inputs are processed by a machine learning model (XGBoost), which analyzes the risk factors and generates an instant prediction. Users receive immediate feedback, such as a high-risk warning prompting them to seek further medical evaluation if needed. Designed for accessibility, the chatbot uses simple Yes/No or numerical responses and includes explanatory notes for medical terms, making it particularly useful in low-resource settings where traditional screening methods may be unavailable. By providing early, non-invasive risk assessment, the chatbot aims to bridge gaps in prenatal care and encourage timely intervention for better maternal and fetal health outcomes. For a clear view refer to appendix 2.

## 4.9 Discussion

Results obtained revealed how machine learning (ML) and digital health technologies can be implemented to overcome the difficulties in detecting gestational diabetes mellitus (GDM), especially in environments with limited resources. For early GDM risk assessment, a scalable, affordable, and easily accessible solution was provided by combining an ML-driven predictive model with a WhatsApp chatbot. The findings were examined below along with their practical ramifications, future directions, and literature review.

### 4.9.1 Performance of Machine Learning Models

To identify the objective of predicting GDM, the study assessed five machine learning algorithms: Random Forest, Logistic Regression, SVM, Decision Tree, and XGBoost. With 96% accuracy, 94% precision, 95% recall, and 94% F1 score, XGBoost was the best-performing model. These findings coincided with prior literature that emphasised the effectiveness of ensemble methods in medical diagnostics (e.g., Zaky et al., 2025; Hassan et al., 2025). Cross-validation further verified XGBoost's robustness, demonstrating its dependability for practical implementation with a mean F1 score of 0.897 with minimal variability.

With metrics similar to XGBoost, Random Forest also demonstrated remarkable performance. This consistency across ensemble methods implies that complex patterns in GDM risk factors can be effectively captured by utilising gradient boosting (as in XGBoost) or combining multiple decision trees (as in Random Forest). However, despite being clinically viable, simpler models such as SVM and Logistic Regression performed slightly lower, most likely because they made linear assumptions and were less able to handle non-linear relationships.

### 4.9.2 Key Risk Factors for GDM

The study found a number of established risk factors for GDM, such as a sedentary lifestyle, age, pregnancy history, family diabetes history, blood pressure and conditions like PCOS or prediabetes ,elevated BMI and abnormal OGTT results . These results highlighted the significance of these variables in clinical screening protocols and support previous research (e.g., Mirabelli et al., 2023; Fortofoiu et al., 2022). In line with their pathophysiological roles in insulin resistance and glucose metabolism, BMI and OGTT levels were found to be the most significant predictors.

The potential for preventive interventions is highlighted by the inclusion of lifestyle factors, such as being sedentary. To reduce the risk of GDM, the model was used, for example, to identify high-risk women early in pregnancy and enable targeted lifestyle changes (e.g., dietary changes, physical activity). This is consistent with the WHO's focus on maternal health from a holistic perspective (WHO, 2023).

### 4.9.3 WhatsApp Chatbot as a Screening Tool

The results of this study showed that, instead of taking the place of crucial prenatal care, the WhatsApp chatbot functioned as an efficient supplementary tool for early GDM risk identification. The chatbot gave women vital health information prior to their first clinical appointment by offering instant preliminary risk assessments. This strategy was especially helpful in low-resource environments where prenatal visits are frequently delayed by obstacles like lack of awareness, transportation difficulties, or financial limitations.

The design of the chatbot effectively bridged the gap between clinical care and home-based screening. Women were able to enter basic health metrics and get immediate feedback regarding their risk level for GDM thanks to its user-friendly interface. The system established a smooth link between initial screening and official healthcare services by producing tailored recommendations for individuals deemed high-risk to pursue additional testing. This feature was particularly useful in areas with poor access to healthcare, where the chatbot could be a crucial triage tool before women could get to hospitals.

The chatbot's conversational, unbiased interface also addressed psychological barriers to early screening. Because of their lack of symptoms or fear of being diagnosed, many women put off or avoid getting tested for GDM. By providing information in an approachable and encouraging way, the chatbot assisted in normalising risk assessment and lowering stigma. Results from pilot testing showed the efficacy of this strategy, with multiple participants stating that after receiving high-risk alerts from the chatbot, they were inspired to seek an earlier clinical evaluation.

Despite its great potential, the chatbot was designed with care to enhance rather than replace in-person clinical visits. The three main purposes of the system were to reduce the number of undiagnosed cases by raising awareness, educate women about the value of early testing, and identify possible high-risk cases before a formal diagnosis. A more complete system that links clinical follow-up and home-based screening could be created in the future by incorporating appointment reminder capabilities and integrating the chatbot with electronic health records. This strategy addresses the unique difficulties of late GDM diagnosis in settings with limited resources while upholding the vital significance of in-person prenatal care.

## 4.10 Conclusion

This chapter presented the results of developing and evaluating machine learning models for predicting Gestational Diabetes Mellitus (GDM) risk, along with the integration of a WhatsApp chatbot for early detection. The findings demonstrated that machine learning algorithms, particularly XGBoost, achieved high predictive performance, with an accuracy of 96%, precision of 94%, recall of 95%, and an F1 score of 94%. Cross-validation further confirmed XGBoost’s robustness, yielding a mean F1 score of 0.897 with low variability, making it the most reliable model for real-world deployment. The Random Forest model also performed strongly, achieving comparable metrics, while Logistic Regression, SVM, and Decision Tree provided moderate yet clinically viable results. The analysis of risk factors reinforced established clinical knowledge, with BMI, OGTT levels, family history of diabetes, and sedentary lifestyle emerging as the most influential predictors of GDM.

The WhatsApp chatbot successfully demonstrated its potential as an accessible, low-cost screening tool, enabling pregnant women in resource-limited settings to assess their GDM risk remotely. The chatbot’s backend, powered by FastAPI, efficiently processed user inputs and delivered real-time predictions, ensuring usability and scalability.

# CHAPTER 5: CONCLUSION AND RECOMMENDATIONS

## 5.1 Introduction

This chapter summarizes the key findings of the study, which aimed to develop a machine learning (ML)-driven predictive model for Gestational Diabetes Mellitus (GDM) risk assessment, integrated with a WhatsApp chatbot for early detection. The study addressed the critical need for accessible and timely GDM screening, particularly in low-resource settings where traditional diagnostic methods are often impractical. The chapter revisits the objectives, evaluates the success of the proposed solutions, and provides recommendations for future research and implementation.

## 5.2 Summary of Findings

The study successfully achieved its primary objective of predicting GDM risk using machine learning models. Five algorithms—Random Forest, Logistic Regression, Support Vector Machine (SVM), Decision Tree, and XGBoost—were evaluated, with XGBoost emerging as the best-performing model based on cross-validation results. The model demonstrated high accuracy (96%), precision (94%), recall (95%), and F1 score (94%), indicating robust performance in identifying GDM cases.

### 5.2.1.The secondary objectives were also met:

Identification of Maternal Risk Factors: Key risk factors such as BMI, OGTT results, family history of diabetes, and sedentary lifestyle were identified as significant predictors of GDM, aligning with existing literature.

.Development of a Predictive Model: The ML models, particularly XGBoost, provided reliable predictions using non-invasive clinical and demographic data.

WhatsApp Chatbot Integration: The chatbot was successfully developed and tested, offering a user-friendly platform for early GDM risk assessment.

## 5.3 Ethical and Practical Considerations

The study complied with ethical standards, which included IRB approval and data anonymisation. However, the model's ability to be validated with Zimbabwean local populations was constrained by its reliance on secondary data. Primary data collection should be given top priority in future research to guarantee cultural and clinical relevance. Furthermore, even though the chatbot protects privacy, continuous oversight is required to protect private health information in accordance with international regulations.

## 5.4 Conclusion

The study underscores the potential of machine learning and digital health tools in improving early detection of GDM. The XGBoost model outperformed other algorithms in cross-validation, achieving the highest mean F1 score (0.897) and demonstrating consistency across validation folds. This highlights its suitability for real-world deployment, where generalizability and reliability are paramount.

The WhatsApp chatbot further enhanced the accessibility of GDM screening, particularly in resource-limited settings. By leveraging widely available technology, the study bridged the gap between advanced ML techniques and practical healthcare delivery. The integration of the chatbot with the predictive model ensured that pregnant women could receive instant risk assessments without the need for invasive tests or multiple clinic visits.

In conclusion, the study successfully addressed its objectives, providing a scalable and effective solution for GDM risk prediction. The combination of XGBoost and the WhatsApp chatbot represents a significant step forward in maternal healthcare, offering a proactive approach to managing GDM and reducing adverse outcomes for mothers and infants.

## 5.5 Recommendations

To build on the findings of this study, the following recommendations are proposed:

Clinical Validation: Future research should include clinical trials to validate the model's performance in diverse populations, ensuring its applicability across different demographic and geographic settings.

Expanded Dataset: Incorporating larger and more diverse datasets could further improve model accuracy and robustness, particularly for underrepresented groups.

Integration with Healthcare Systems: Collaborations with healthcare providers and policymakers could facilitate the integration of the chatbot into routine antenatal care, ensuring widespread adoption.

Longitudinal Studies: Tracking the long-term health outcomes of women who use the chatbot could provide insights into its impact on GDM management and prevention of type 2 diabetes post-pregnancy.

By addressing these areas, future iterations of this project can further enhance its effectiveness and contribute to global efforts in combating GDM and improving maternal health outcomes.

This chapter concludes the dissertation, summarizing the achievements, highlighting the success of the XGBoost model, and outlining actionable recommendations for future work. The study’s contributions pave the way for innovative, technology-driven solutions in maternal healthcare.

# **REFERENCES**

**A. Damayanti and A. Baita** “Comparison of Support Vector Machine (SVM) and Random Forest (RF) Algorithm Performance with Random Undersampling Technique to Predict Gestational Diabetes Mellitus Risk” [Journal]. - [s.l.] : JAIC, 2025.. - 2 : Vol. 9.

**ACOG.** Gestational Diabetes Mellitus. [Report]. - [s.l.] : Practice Bulletin No. 190., 2023.

**Afrizal S. H., Hakiem, N., Permanasari, A. E., Albab, H. S., Sanjaya, G. Y., & Lazuardi, L.** A User-Centered Design of Natural Language Processing for Maternal Monitoring Chatbot System. [Journal]. - 2022.

**Artzi N.S., et al.** EHR-based prediction of gestational diabetes. [Journal]. - [s.l.] : Nature Medicine , 2020. - 1 : Vol. 26.

**Association American Diabetes** Standards of Medical Care in Diabetes—2023. Diabetes Care [Journal]. - [s.l.] : Diabetes Care , 2023. - Supplement1 : Vol. 46.

**Belsti Y., Moran, L., Du, L., Mousa, A., De Silva, K., Enticott, J., & Teede, H. J.** Comparison of machine learning and conventional logistic regression-based prediction models for gestational diabetes in an ethnically diverse population; The Monash GDM Machine Learning Model. [Journal]. - [s.l.] : International Journal of Medical Informatics., 2023.

**Bentéjac C, Csörgő, A & Martínez-Muñoz, G** A comparative analysis of XGBoost’, Artificial Intelligence Review, [Article]. - 2021. - 3 : Vol. 54.

**Bhaskar P., Kunchala, T., Peddi, S., Syed, R., & Namepalli, P. (2024).** A Chatbot System for Supporting Women and Families during Pregnancy. [Journal]. - [s.l.] : International Journal of Innovative Science and Research Technology., 2024.

**Bigdeli S. K., Ghazisaedi, M., Ayyoubzadeh, S. M., Hantoushzadeh, S., & Ahmadi, M.** Predicting Gestational Diabetes Mellitus in the first trimester using machine learning algorithms: a cross-sectional study at a hospital fertility health center in Iran. [Journal]. - [s.l.] : BMC Medical Informatics and Decision Making, 2025. - 1 : Vol. 25.

**Bishop C. M.** Pattern Recognition and Machine Learning [Article]. - [s.l.] : Springer., 2006.

**Breiman L.** Random forests. Machine Learning [Article]. - 2001. - 1 : Vol. 45.

**Cascabulho A. M. E., Alves, C. V. de L., Neto, D. G., Antunes, G. P. D., Bastos, J., Balthar Achi Boechat, J. H., Beltrão, L. P. V., Souza, M. B. de, Carmo, V. S. T. do, & Magalhães, W. L. R.** Diabetes mellitus gestacional: fatores de risco, complicações e tratamento [Journal]. - 2024.

**Chen P, Wang S, Ji J, Ge A, Chen C, Zhu Y, Xie N, Wang Y.**, 2015. Risk factors and management of gestational diabetes.. Cell Biochem Biophys., 71(2), pp. 689-694.

**Chen T & Guestrin, C** XGBoost: a scalable tree boosting system’, Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining [Article]. - 2016.

**Cortes C., & Vapnik, V.** Support-vector networks. Machine Learning [Article]. - 1995. - 3 : Vol. 20.

**Draelos Rachel** Measuring Performance: The Confusion Matrix [Article]. - 2019.

**Fortofoiu M.** Analysis of Risk Factors for the Development of Gestational Diabetes Mellitus in a Group of Romanian Patients. [Journal]. - 2022.

**Gwarisa Mitchell** Hypertension Emerges as Leading Killer of Pregnant Women in Zimbabwe [Report]. - [s.l.] : HealthTimes, 2025.

**Hassan A Ahmad S ,Iqbal T, Munir E U,Ayyub K and Ramzan N** Enhanced Model for Gestational Diabetes Mellitus Prediction Using a Fusion Technique of Multiple Algorithms with Explainability [Journal]. - [s.l.] : Int J Comput Intell Syst , 2025. - 47 : Vol. 18.

**Hirst JE, R.-G. C. a. J. H.**, 2012. A systematic review of trends of gestational diabetes mellitus in Asia. Journal of Diabetology. , 3(4).

**Hosmer D. W., & Lemeshow, S.** Applied Logistic Regression [Article]. - Wiley : [s.n.], 2000. - 2nd.

**Iman A.E.H. [et al.]** Prevalence and Risk Factors of Gestational Diabetes Mellitus in Romania: Maternal and Fetal Outcomes. [Journal]. - [s.l.] : Medicina , 2025. - 2 : Vol. 61.

**International Diabetes Association** ., 2009. Standards of medical care in diabetes.. Diabetes Care., 32(Suppl 1), p. S13–S61.

**International Diabetes Association**, https://idf.org/about-diabetes/gestational-diabetes/ 2024.

**James G, Witten, D, Hastie, T & Tibshirani, R ,** An introduction to statistical learning [Article]. - [s.l.] : Springer, 2021. - 2nd edn.

**Kansu-Celik,Ozgu-EArdinc A.S, Kisa-Karakaya, Tasci Y and Erkaya S.**, 2019. Fasting and postprandialmplasma glucose screening for gestational diabetes mellitus. World Health Organisation , 25(4), pp. 282-289.

**Kaya Y Bütün Z,Çelik, Ö. Salik E A ,Tahta T , Yavuz A A** The early prediction of gestational diabetes mellitus by machine learning models. [Journal]. - [s.l.] : BMC Pregnancy Childbirth , 2024. - 574 : Vol. 24.

**Li G., et al.** Machine learning for GDM prediction. [Journal]. - [s.l.] : Journal of Medical Internet Research, 2023. - 1 : Vol. 25.

**Liaw A., & Wiener, M.** Classification and regression by randomForest. R News [Article]. - 2002. - 3 : Vol. 2.

**Loh W-Y** Fifty years of classification and regression trees’ [Article]. - [s.l.] : International Statistical Review,, 2014. - Vol. 82.

**McIntyre H. D., Catalano, P., Zhang, C., Desoye, G., Mathiesen, E. R., & Damm, P.** Gestational diabetes mellitus. [Journal]. - [s.l.] : Nature Reviews Disease Primers, 2022. - 1 : Vol. 8.

**McIntyre H. D., et al.** Gestational diabetes mellitus. [Journal]. - [s.l.] : Nature Reviews Disease Primers, 2022. - 1 : Vol. 8.

**Menard S.** Applied Logistic Regression Analysis [Article]. - Sage : [s.n.], 2002. - 2nd ed.

**Metzger B. E., et al.** Hyperglycemia and adverse pregnancy outcomes [Journal]. - [s.l.] : NEJM, 2008. - 19 : Vol. 358.

**Mirabelli M., Tocci, V., Donnici, A., Giuliano, S., Sarnelli, P., Salatino, A., Greco, M., Puccio, L., Chiefari, E., Foti, D., & Brunetti, A.** Maternal Preconception Body Mass Index Overtakes Age as a Risk Factor for Gestational Diabetes Mellitus. [Journal]. - 2023. - 8 : Vol. 12.

**Molinaro AM, Dudoit, S & van der Laan, MJ** Tree-based methods’, in Handbook of statistical genomics [Article]. - Wiley : [s.n.], 2021. - 4th edn.

**Mugoye K., Okoyo, H., & Mcoyowo, S.** Smart-bot Technology: Conversational Agents Role in Maternal Healthcare Support. IST-Africa Week Conference [Journal]. - 2019.

**Nhidza G, Mutsaka K, Malunga G and Zhou D** 2018. Diagnosis of Gestational Diabetes Mellitus in Urban Harare, Zimbabwe.. The Open Public Health Journal., Volume 11, pp. 1-7.

**Niharika P Sai K,Neeraja K ,Santhadevi D and Renugadevi N** Predictive framework for Gestational Diabetes Mellitus using Machine learning models and deep learning techniques [Journal]. - 2024.

**Organization World Health** Digital interventions for health systems. [Journal]. - 2021.

**Probst P, Wright, MN & Boulesteix, A-L** Hyperparameters and tuning strategies for random forest’, Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery [Article]. - 2019. - 3 : Vol. 9.

**Puspitasari I. W., Rinawan, F. R., Purnama, W. G., Susiarno, H., & Susanti, A. I.** Development of a Chatbot for Pregnant Women on a Posyandu Application in Indonesia: From Qualitative Approach to Decision Tree Method. [Journal]. - [s.l.] : Informatics (Basel), 2022. - 4 : Vol. 9.

**Sacks D.B., et al.** Guidelines for laboratory diagnosis of diabetes. Diabetes Care [Journal]. - 2022. - S1 : Vol. 45.

**Sagstad M. H., Morken, N.-H., Lund, A., Dingsør, L. J., Nilsen, A. B. V., & Sørbye, L. M.** Quantitative User Data From a Chatbot Developed for Women With Gestational Diabetes Mellitus: Observational Study. [Journal]. - [s.l.] : JMIR Formative Research, 2022. - 4 : Vol. 6.

**Tiwari S., Gochhait, S., & Bonuga, H.** Exploring Machine Learning Classifiers for Identifying Gestational Diabetes Mellitus. [Journal]. - 2024.

**Wang X., Wang, Y., Zhang, S., Yao, L., & Xu, S.** Analysis and Prediction of Gestational Diabetes Mellitus by the Ensemble Learning Method. [Journal]. - [s.l.] : International Journal of Computational Intelligence Systems, 2022. - 1 : Vol. 15.

**WHO.** WHO recommendations on antenatal care for a positive pregnancy experience [Journal]. - 2023.

**World Health Organisation** , https://www.who.int/news-room/fact-sheets/detail/diabetes 2024.

**Zaky H Fthenou e,Srour L,Farrell T,Bashir M,Hajj N and Alam T** Machine learning based model for the early detection of Gestational Diabetes Mellitus [Journal]. - [s.l.] : BMC Med Inform Decis Mak, 2025. - 130 : Vol. 25.

**Zhou W., et al.** Barriers to gestational diabetes screening in low-resource settings. [Journal]. - [s.l.] : Lancet Global Health, 2021. - 3 : Vol. 9.

**Zhou Z-H** Ensemble methods: foundations and algorithms, [Article]. - Boca Raton. : CRC Press, 2021. - 2nd edn,.

**Басараб М. Р., Іванько, Е. О., & Кулкарні, В.** Prediction of the Development of Gestational Diabetes Mellitus in Pregnant Women Using Machine Learning Methods. [Journal]. - 2021.

# APPENDICES

## APPENDIX 1: JREC APPROVAL FORM

## C:\Users\Shamiso\Downloads\IMG_7791.PNGC:\Users\Shamiso\Downloads\IMG_7790.PNGAPPENDIX 2: Chart Bot Output

