A Project report on

PREDICTING DIABETES IN PREGNANT WOMANS AND NEONATAL MELLITUS IN NEW BORN CHILD USING MACHINE LEARNING

A Dissertation submitted to JNTU Hyderabad in partial fulfillment of the Academic requirements for the award of the degree.

Bachelor of Technology

in

Computer Science and Engineering

Submitted by

G.Mahendra
(21H55A0504)
H.Laxman
(21H55A05058)
M.Jyoshna
(21H55A0513)

Under the esteemed guidance of
Mr. J. Ranjith
(**Assistant Professor**)

Department of Computer Science and Engineering

CMR COLLEGE OF ENGINEERING & TECHNOLOGY
(UGC Autonomous)
*Approved by AICTE *Affiliated to JNTUH *NAAC Accredited with A^+ Grade
KANDLAKOYA, MEDCHAL ROAD, HYDERABAD - 501401.

2020- 2024

CMR COLLEGE OF ENGINEERING & TECHNOLOGY

KANDLAKOYA, MEDCHAL ROAD, HYDERABAD – 501401

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the Major Project Phase I report entitled **"Predicting Diabetes in pregnant women's and neonatal mellitus in new born child using Machine Learning** being submitted by G.Mahendra(21H55A0504), H.Laxman (21H55A0508), M.Jyoshna(21H55A0513) in partial fulfillment for the award of **Bachelor of Technology in Computer Science and Engineering** is a record of bonafide work carried out his/her under my guidance and supervision.

The results embodies in this project report have not been submitted to any other University or Institute for the award of any Degree.

Mr. J. Ranjith
Assistant Professor
Dept. of CSE

Dr. Siva Skandha Sanagala
Associate Professor and HOD
Dept. of CSE

ACKNOWLEDGEMENT

With great pleasure we want to take this opportunity to express my heartfelt gratitude to all the people who helped in making this project work a grand success.

We are grateful to **Mr. J. Ranjith, Assistant Professor**, Department of Computer Science and Engineering for his valuable technical suggestions and guidance during the execution of this project work.

We would like to thank **Dr. Siva Skandha Sanagala,** Head of the Department of Computer Science and Engineering, CMR College of Engineering and Technology, who is the major driving forces to complete my project work successfully.

We are very grateful to **Dr. Vijaya Kumar Koppula**, Dean-Academics, CMR College of Engineering and Technology, for his constant support and motivation in carrying out the project work successfully.

We are highly indebted to **Major Dr. V A Narayana,** Principal, CMR College of Engineering and Technology, for giving permission to carry out this project in a successful and fruitful way.

We would like to thank the **Teaching & Non- teaching** staff of Department of Computer Science and Engineering for their co-operation

We express our sincere thanks to **Shri. Ch. Gopal Reddy**, Secretary, CMR Group of Institutions, for his continuous care.

Finally, We extend thanks to our parents who stood behind us at different stages of this Project. We sincerely acknowledge and thank all those who gave support directly and indirectly in completion of this project work.

G.Mahendra 21H55A0504
H.Laxman 21H55A0508
M.Jyoshna 21H55A0513

INDEX

CHAPTER NO	TITLE	PAGE NO
	LIST OF FIGURES	14
	LIST OF TABLES	5
	ABSTRACT	3
1	INTRODUCTION	4-7
	1.1 problem statement	7
	1.2 research objectives	8
	1.3 future scope and limitations	8
2	BACKGROUND WORK	9
	2.1 < existing method 1 >	10-12
	2.1.1 introduction	12-13
	2.1.2 merits ,demerits challenges	13
	2.1.3 implementation	13-14
	2.2 < existing method 2 >	15
	2.2.1 introduction	15
	2.2.2 merits and demerits	15
	2.2.3 implementation	16
3	RESULTS AND DISCUSSION	17
	3.1 Performance metrics	18-23
4	CONCLUSION	24
	4.1 conclusion	25
	REFERENCE	26
	reference	27

LIST OF FIGURES

FIGURE NO	TITLE	PAGE NO
1.1	architecture	6
1.2	Data preprocessing diagram	6
1.3	Model planning diagram	7
3.1	Input slide-1	18
3.2	Input slide-2	19
3.3	Training data stats 1	19
3.4	Training data stats 2	20
3.5	Patient data	20
3.6	Pregnancy count graph	20
3.7	Glucose value graph	21
3.8	Skin thickness value count	21
3.9	Insulin value count	22
3.10	BMI value graph	22
3.11	DPF value count	23

ABSTRACT

Diabetes during pregnancy is a major source of health problems in unborn infants and their moms. Because gestational diabetes can develop to permanent diabetes, ML is an important method for predicting the likelihood of such progression based on the given features. Although the current study may predict lifelong diabetes in pregnant women, it cannot predict the likelihood of neonatal diabetes. As a result, new characteristics are required to improve the forecasting of neonatal mellitus and provide the most accurate diabetes persistence results in pregnant women. Python scripting and the application of Machine Learning methods such as SVM, KNN, and LR can assist in achieving this aim. The preprocessing ML dataset focusing on Diabetes from the Pima Indian diabetes database collected through Kaggle. In addition, two new attributes were added to the project's dataset. According to research, machine learning models using characteristics like SVM and decision trees may successfully predict the risk of diabetes in pregnant women. Various factors have been used to predict the beginning of this condition during pregnancy.

CHAPTER 1
INTRODUCTION

CHAPTER I

INTRODUCTION

learning algorithms utilize previous data to discover patterns and make predictions about new data. In the case of diabetes prediction, machine learning algorithms may use data from Machine electronic health records, clinical notes, test findings, and patient demographics to identify people at risk of acquiring diabetes. Several research have looked at the use of ML algorithms to predict diabetes in pregnant women and neonatal diabetes in babies. One research, for example, employed ML algorithms to fore the risk of gestational diabetes in pregnant women based on demographic and clinical variables. In terms of predicting gestational diabetes, the algorithm has an accuracy of 89.4%. disease that can occur throughout pregnancy. Because of this complexity, developing reliable prediction models that account for all of these elements and changes is difficult. These algorithms make use of a many of data sources, such as Clinical factors, Laboratory tests, and Demographic data. Using these data sets, the algorithms can accurately identify women who are at high risk of having GDM.NDM (neonatal diabetes mellitus) is an uncommon type of diabetes that develops within the first six months of life. Genetic mutations that impair insulin synthesis or function create the disease. NDM is a diverse disorder with different genetic alterations in different people. NDM can be temporary, resolving within a few months of birth, or persistent, necessitating lifetime car Gathering relevant data from pregnant women and newborns, including medical history, glucose levels, BMI, and other factors. This data is preprocessed, selecting important features and ensuring data quality. ML algorithms such as LR, DT, or NN are then trained using this data to establish patterns and relationships between the features and diabetes outcomes. The trained model is evaluated using validation data, and once deemed satisfactory, it can be utilized to predict diabetes risk in pregnant women and neonatal diabetes in newborns. These predictions aid in identifying high-risk cases and enabling timely interventions and monitoring by healthcare professionals

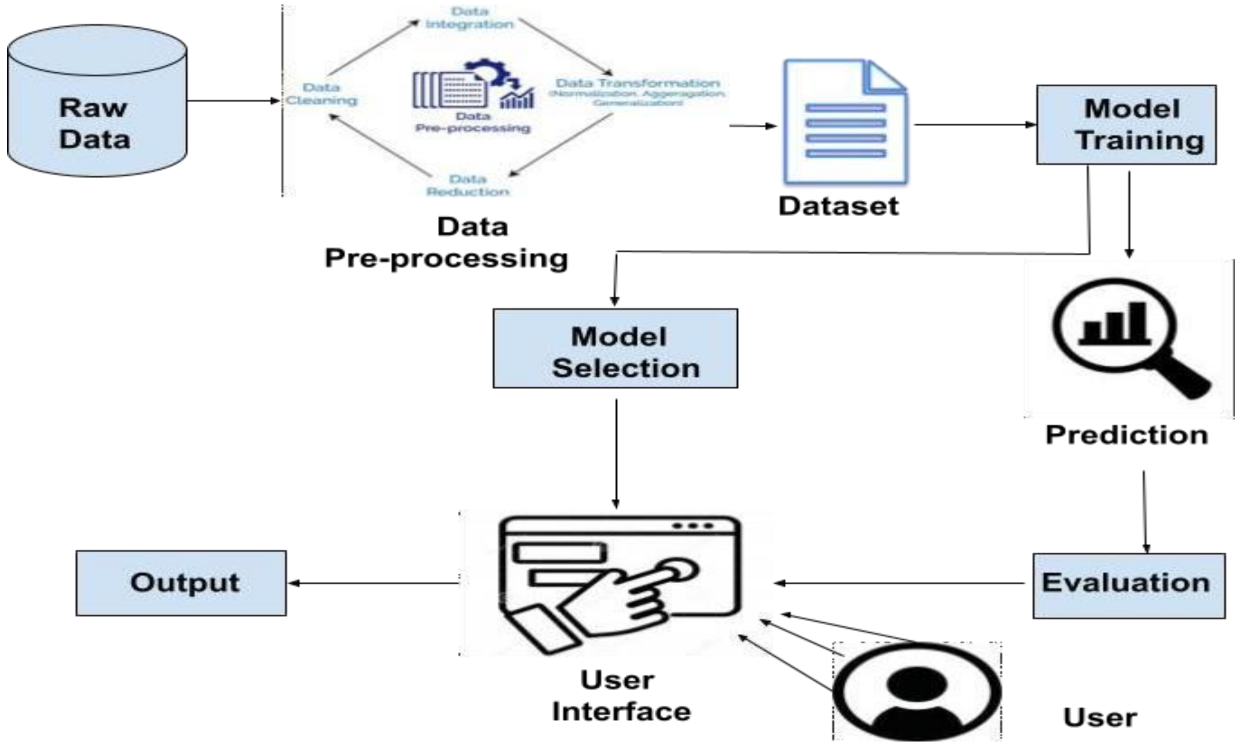

Figure 1.1 architecture

1. **Raw Dataset** The Raw Dataset serves as the starting point for the architecture. This dataset is specifically curated to contain a collection of data of various females with symptoms of diabetes. The data is typically sourced from various real-world scenarios or collected under specific conditions.
2. **Data Preprocessing** Data preprocessing is an essential step in preparing the blurred images from the Real Blur dataset for training and evaluation. In this step, several operations are performed to transform the raw image data into a suitable format for feeding into a neural network. The main operations typically include conversion to tensors and resizing

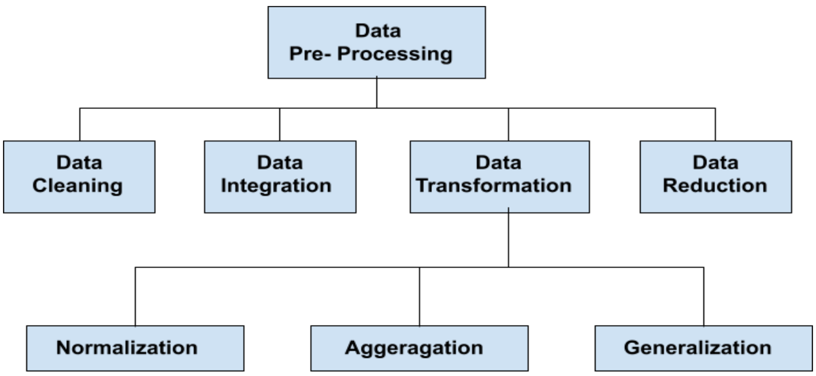

Figure 1.2. Data Preprocessing diagram

3. **Model Training** Model Training is a Crucial step in the project as the various models like Naive Bayes, K-Nearest Neighbour, Random Forest, Logistic Regression, SVM are used to train the dataset and the best of the mentioned is used for final execution.
4. **Prediction** Predictions are drawn from the training process to know that which model is better in the prediction.
5. **Model Evaluation** In this step, the model is evaluated using the preprocessed testing data, which consist of diabetes data of females and hence the best model is drawn out of the five.
6. **User Interface** Here the user can input the data he has and then the algorithm runs in the background to calculate whether the user has diabetes or not.
7. **Output** Here the algorithm generates output and displays prompts either "You are diabetic" or "You are not diabetic". Along with the accuracy of the calculation.

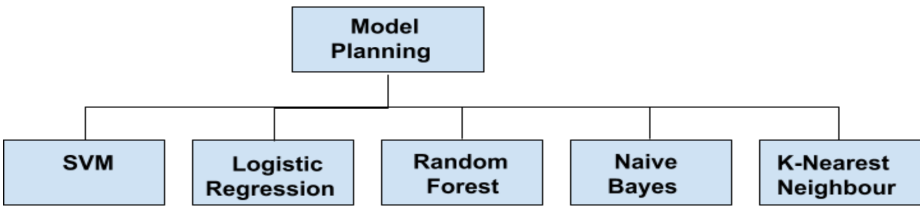

Figure 1.3. Model Planning Diagram

1.1 Problem statement

Diabetes is a chronic condition that affects millions of individuals all over the world and can have major health repercussions, particularly during pregnancy. Gestational diabetes mellitus (GDM) is a frequent condition that can result in neonatal mellitus in neonates. Early identification and treatment of GDM are critical for avoiding negative effects for both the mother and the child. Current GDM detection approaches rely heavily on blood glucose tests and clinical risk factors. However, these strategies may not be precise enough to forecast GDM in all circumstances. Machine learning (ML) has showed potential in predicting GDM by analysing massive datasets and finding risk variables that traditional approaches may miss. The problem statement for utilising machine learning to predict diabetes in pregnant women and neonatal mellitus in newborns entails establishing accurate and reliable ML models that can predict the GDM in pregnant women and neonatal mellitus .

1.2 Research Objective

The project's goal is to use machine learning techniques to forecast the risk of permanent diabetes in pregnant women and Neonatal Mellitus in newborn children. By developing machine learning models and providing a user interface for on-spot data calculation, the project seeks to facilitate early detection and prevention of diabetes in these populations. Large datasets will be analysed using machine learning algorithms to uncover patterns and risk variables related with persistent diabetes in pregnant women and Neonatal Mellitus diabetes in newborns. By training these models on comprehensive and diverse datasets, the goal is to ensure their reliability and generalizability in real-world scenarios.

1.3 Future scope and limitations

The potential of machine learning to predict gestational diabetes and neonatal mellitus is promising and encompasses several potential advancements and areas of exploration. Here are some key aspects that offer significant opportunities for further development.
Accuracy- Researchers can focus on refining existing machine learning models and developing more advanced algorithms to improve the accuracy of predictions. This involves incorporating additional data sources, such as genetic information or nove biomarkers, tenhance the predictive power of the models. Longitudinal Analysis-The integration of longitudinal data, including multiple measurements taken throughout pregnancy, can provide a more comprehensive understanding of the dynamic changes in risk factors and their impact on diabetes development .Personalized Risk Stratification- The future lies in developing personalized risk stratification models that can consider individual characteristics, including genetic predisposition, socio-demographic factors, and lifestyle choices. By tailoring predictions to specific individuals, healthcare providers can offer personalized interventions and support, leading to more effective prevention and management strategies

CHAPTER 2
BACKGROUND WORK

CHAPTER 2 BACKGROUND

WORK

2.1 EXISTING METHOD 1:

Summary of Existing Approaches

Reference number	Author(s)	Paper Name	Algorithm Used	Results/Remarks
[1]	AV Srinivas, Abbireddy Ramya, G.Chandralekha	"Prediction of diabetes using Machine Learning" (2022)	Random Forest, SVM, Logistic Regression.	Random Forest provided the highest accuracy with 92% compared to other models.
[2]	Quan Zou , Kaiyang Qu , Yamei Luo	**"**Predicting Diabetes Mellitus With Machine Learning Techniques**"** **(**2018**)**	Random forest , Decision tree , Neural network	The results showed that prediction with random forest could reach the highest accuracy (ACC = 0.8084) when all the attributes were used.
[3]	Марко Романович Басараб,Екатерина Олеговна Іванько	"Prediction of the Development of Gestational Diabetes Mellitus in Pregnant Women Using Machine Learning Methods"	SVM	Support vector machine classifier demonstrated thehighestaccuracy,78.1% predicting thedevelopment of gestational diabetes based.

[4]	Zahura Zaman , Md. Ashrak Al Arif Shohas, Mahedi Hasan ,Meherab Hossain	"Assessing Machine Learning Methods for Predicting Diabetes among Pregnant Women."	SVM, Decision tree, Naive Bayes.	The most accurate method was determined to be logistic regression and this resulted in an accuracy of 78.01 percent.
		(2022)		

[5]	Lili Wei , Yueshuai Pan , Yan Zhang	“Application of machine learning algorithm for predicting gestational diabetes mellitus in early pregnancy” (2021)	Random Forest	The overall predictive accuracy of the F1 model was 93.10%,the predictive accuracy of GDM-positive cases 37.10%. The corresponding values for the F2 model 88.70%, and 79.44%.
[6]	Thach Tran , Jane E.Hirst	“Early prediction of Gestational Diabetes Mellitus in Vietnam” (2013)	The ADA prognostic model, consisting of age and BMI at booking	Selective screening of women for GDM using the ADA model with a risk threshold of 3% gave 93% sensitivity for identification of women with GDM with a 27% reduction in the number of OGTTs required.
[7]	Sara L. White , Debbie A Lawlor , Annette Briley	“Early Antenatal Prediction of Gestational Diabetes in Obese Women” (2016)	stepwise logistic regression	(age, previous GDM, family history of type 2 diabetes, systolic blood pressure, sum of skinfold thicknesses, waist:height and neck:thigh ratios) provided an area under the curve of 0.71 (95%CI 0.68–0.74)
[9]	M Sekulić , D Relić	“The role of machine learning in identification of early gestational diabetes mellitus prediction “	Machine learning	microRNA and visceral fat mass as novel biomarkers GDMprediction.recommended division of data into training(2/3) and validation(1/3) datasets

[9]	M Sekulić , D Relić	“The role of machine learning in identification of early gestational diabetes mellitus prediction “	Machine learning	microRNA and visceral fat mass as novel biomarkers GDMprediction.recommended division of data into training(2/3) and validation(1/3) datasets
[10]	XINXI LU , JIKAI WANG	“Prediction of Gestational Diabetes and Hypertension using	Logistic Regression, XGBoost, LightGBM, and Neural Network	The predictive AUCs for diabetes and hypertension can reach 0.92 and 0.87

		Eximanation data” (2022)	Model Logistic Regression, XGBoost, LightGBM, and Neural Network Model	
[11]	Jie Zhang , Fang Wang	“Prediction of Gestational Diabetes Mellitus under Cascade and Ensemble Learning Algorithm” (2022)	Logistic Regression	On the data set utilized in this work, the accuracy of the proposed prediction model is 80.3%, the precision is 74.6%, and the recall rate is 79.3%.
[12]	Jenny Yang , David Clifton	**“**Machine LearningBased Risk Stratification for Gestational Diabetes Management**”** (2022)	linear and non-linear tree-based regression models, Random forest	Random Forest model (0.022 [0.021–0.024], 0.447 [0.400–0.482], and 0.117 [0.117 (0.114–0.121], for MSE, R2, MAE, respectively)

2.1.1 INTRODUCTION

Diabetes is one of the most hazardous diseases on the planet. Diabetes is caused by obesity, excessive blood glucose levels, and other factors. It alters the hormone insulin, causing aberrant crab metabolism and improving blood sugar levels. When the body does not produce enough insulin, diabetes develops. According to the World Health Organization, 422 million people worldwide suffer from diabetes, with the majority living in low- or middle-income nations. Up until 2030, this figure might be boosted to 490 billion. Diabetes is, nevertheless, prevalent in a number of countries, including Canada, China, and India. With a population of more than 100 million people, India has a total of 40 million diabetes. While we didn't achieve our goal of 100 percent accuracy in diabetes prediction, we did develop a system that can come close to it given enough time and data. As with any project of this nature, there is room for improvement. Because of the nature of this project, multiple algorithms can be combined as modules and their results combined to improve the accuracy of the final result. This research could be expanded to see how likely non-diabetic people are to develop diabetes in the YMER || ISSN : 0044-0477 VOLUME 21 : ISSUE 5 (May) - 2022 http://ymerdigital.com Page No:485 coming years. Thus, for this purpose we apply popular classification and ensemble methods on dataset for prediction. Diabetes is a common chronic disease that threat to human health. As a result, we use common classification and algorithms on the dataset to make predictions. Diabetes is a prevalent chronic disease that can be extremely dangerous to one's health. Diabetes is diagnosed when blood glucose levels are greater than normal, which is caused by insulin secretion or biological factors. Diabetes can harm our bodies in a variety of ways, including causing tissue, kidney, eye, and blood artery dysfunction. Based on physical examination data and consultation with doctors, machine learning may make a preliminary diagnosis of diabetes mellitus. Many techniques, including machine learning methods like Random Forest, Support Vector Machine, Decision Tree, and others, have recently been utilised to predict diabetes. We can forecast diabetes using machine learning approaches by creating predicting models based on medical datasets.

2.1.2 Merits

1. Early detection allows timely interventions and improves outcomes.

2. Personalized care plans based on individual risk factors enhance treatment effectiveness.

3. Risk stratification optimizes healthcare delivery and resource allocation.

4. Cost savings through reduced complications and hospitalizations.

5. Insights gained contribute to further research and understanding of diabetes.

2.1.3 Demerits

• The assessment is still done manually, resulting in many

2.1.4 Challenges

While the application of ML algorithms for foreseeing diabetes in pregnant women and neonatal diabetes mellitus shows considerable potential, there are various issues that must be solved in order for these models to be accurate and useful. One of the most difficult aspects of forecasting diabetes in pregnant women is the condition's heterogeneity. Gestational diabetes mellitus (GDM) is a complicated illness impacted by a variety of variables such as maternal age, BMI, ethnicity, and diabetes family history. Furthermore, GDM is a changing

2.1.5 implementation

Gradient Boosting – Gradient Boosting is most powerful ensemble technique used for prediction and it is a classification technique. It combine week learner together to make strong learner models for prediction. It uses Decision Tree model. it classify complex data sets and it is very effective and popular method. In gradient boosting model performance improve over iterations.

Algorithm-

>Consider a sample of target values as P >Estimate

the error in target values.

>Update and adjust the weights to reduce error M.

>P[x] =p[x] +alpha M[x]

>Model Learners are analyzed and calculated by loss function F >Repeat

steps till desired & target result P.

MODEL BUILDING

This is most important phase which includes model building for prediction of diabetes. In this we have implemented various machine learning algorithms which are discussed above for diabetes prediction.

Procedure of Proposed Methodology-

Step1: Import required libraries, Import diabetes dataset.

Step2: Pre-process data to remove missing data.

Step3: Perform percentage split of 80% to divide dataset as Training set and 20% to Test set.

Step4: Select the machine learning algorithm i.e. K- Nearest Neighbour, Support Vector Machine, Decision Tree, Logistic regression, Random Forest and GradiboostialgorithStep5:

2.2 EXISTING METHOD 2

2.2.1 INTRODUCTION

Diabetes is the fast growing disease among the people even among the youngster. In understanding diabetes and how it develops, we need to understand what happens in the body without diabetes. Sugar (glucose) comes from the foods that we eat, specifically carbohydrate foods. Carbohydrate foods provide our body with its main energy source everybody, even those people with diabetes, needs carbohydrate. Carbohydrate foods include bread, cereal, pasta, rice, fruit, dairy product and vegetables (especially starchyvegetables). When we eat these foods, the body breaks them down into glucose The glucose moves around the body in the blood stream. Some of the glucose is taken to our brain to help us think clearly and function. The remainder of the glucose is taken to the cells of our body for energy and also to our liver, where it is stored as energy that is used later by the body. In order for the body to use glucose for energy, insulin is required. Insulin is a hormone that is produced by the beta cells in the pancreas. Insulin works like a key to a door. Insulin attaches itself to doors on the cell, opening the door to allow glucose to move from the blood stream, through the door, and into the cell If the pancreas is not able to produce enough insulin(insulin deficiency) or if the body cannot use the insulin it produces (insulin resistance), glucose builds up in the bloodstream (hyperglycaemia) and diabetes develops. Diabetes Mellitus means high levels of sugar (glucose) in the blood stream and in the urine

2.2.2 merits

Higher accuracy can be achieved with less training data.

Use of a Machine Learning Algorithm improves the system's reliability and accuracy.

2.2.3demerits

The assessment is still done manually, resulting in many inaccuracies and incorrect treatments

2.2.4 Implementation

The prediction of diabetes in pregnant women and neonatal mellitus using machine learning has several future enhancements that can further improve the accuracy and effectiveness of the predictive models. Here are some key areas for future development.

Incorporating multi-omics data, such as genomes, transcriptomics, proteomics, and metabolomics, can give a more thorough picture of the molecular mechanisms underlying gestational diabetes and neonatal mellitus. Integrating these multi-dimensional datasets with machine learning models can uncover novel biomarkers, identify disease pathways, and enhance the accuracy of predictions.

Real-Time Data Incorporation- The incorporation of real-time data, such as continuous glucose monitoring, maternal vital signs, and foetal monitoring, can enable dynamic and personalized risk assessment. By incorporating real-time data streams into machine learning models, healthcare providers can receive timely alerts and take proactive measures to prevent complications associated with gestational diabetes and neonatal mellitus.

Integration of Electronic Health Records (EHR)- Taking use of the large quantity of information included in electronic health records can provide a comprehensive patient profile for accurate risk assessment. By integrating EHR data with machine learning models, healthcare providers can tap into historical medical records, previous pregnancies, medication history, and comorbidities to improve prediction accuracy and develop personalized interventions.

CHAPTER 3

RESULT AND DISCUSSION

Result

The diabetes dataset consists of various attributes that determine a person's risk of being diagnosed by diabetes. It contains data of various patients bearing one or more of the attributes mentioned in the data set. The detailed meanings of each attribute mentioned in the taken dataset is as follows.

Glucose Level- The concentration of glucose (sugar) in the patient's blood is represented by this aspect. Milligrams per deciliter (mg/dL) or millimoles per litre (mmol/L) are the most common units of measurement. Elevated glucose levels might be a sign of poor glucose tolerance or diabetes.

Blood Pressure- Typically, this characteristic comprises two values: systolic and diastolic blood pressure. When the heart contracts, systolic blood pressure indicates the pressure in the arteries, whereas diastolic blood pressure shows the pressure when the heart is at rest. Hypertension (high blood pressure) is a risk factor for gestational diabetes.

BMI (Body Mass Index)- BMI is a measure of body fat that is dependent on a person's height and weight. It is computed by dividing the weight in kilograms by the height in meters squared. BMI indicates if a patient is underweight, normal weight, overweight, or obese. Obesity is a major risk factor for developing gestational diabetes.

Performance metrics

Figure 3.1 input slide 1

Figure 3.2 . Input Slider-2

Figure 3.3. Training Data Stats 1

	Pregnancies	Glucose	BloodPressure	SkinThickness
count	768	768	768	768
mean	3.8451	120.8945	69.1055	20.5365
std	3.3696	31.9726	19.3558	15.9522
min	0	0	0	0
25%	1	99	62	0
50%	3	117	72	23
75%	6	140.25	80	32
max	17	199	122	99

	se	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
count	68	768	768	768	768	768	768	768
mean	45	69.1055	20.5365	79.7995	31.9926	0.4719	33.2409	0.349
std	26	19.3558	15.9522	115.244	7.8842	0.3313	11.7602	0.477
min	0	0	0	0	0	0.078	21	0
25%	99	62	0	0	27.3	0.2438	24	0
50%	17	72	23	30.5	32	0.3725	29	0
75%	25	80	32	127.25	36.6	0.6263	41	1
max	99	122	99	846	67.1	2.42	81	1

Figure 3.4 Training Data Stats-2

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age
0	3	120	70	20	79	20	0.47	33

Figure 3.5 Patient Data

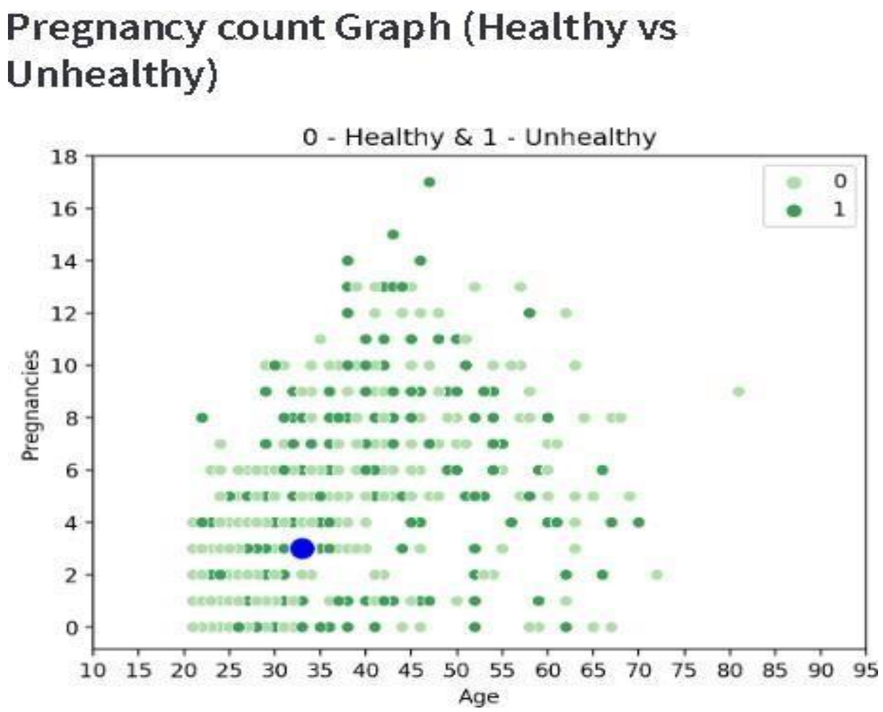

Figure 3.6 pregnenecy count graph

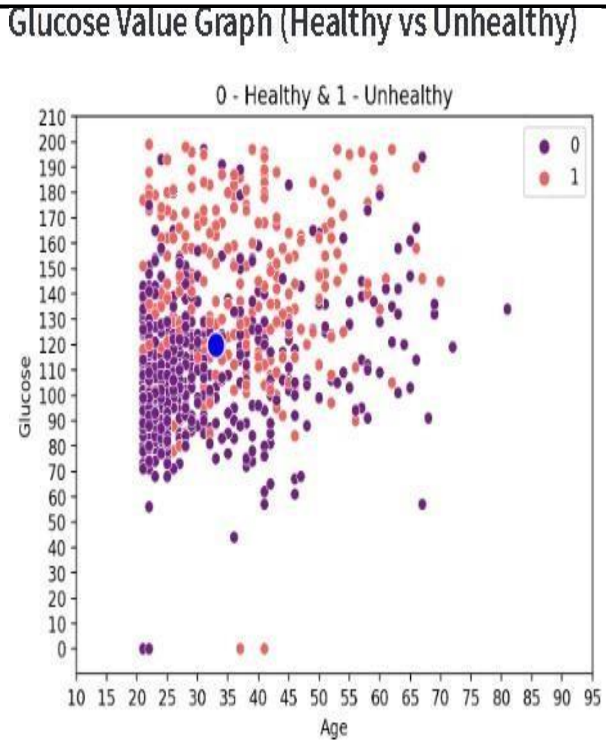

Figure 3.7 Glucose Value Graph

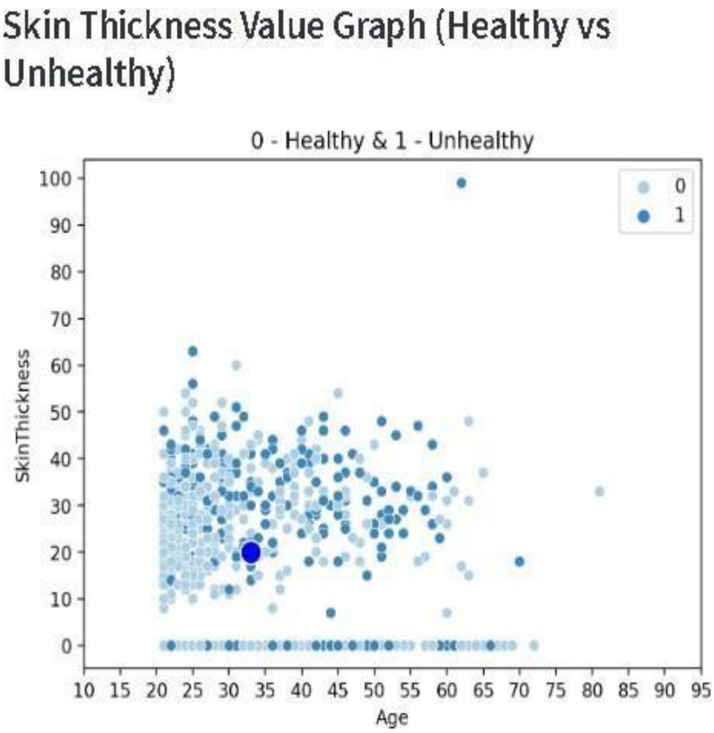

Figure 3.8 skin thickness value ciunt

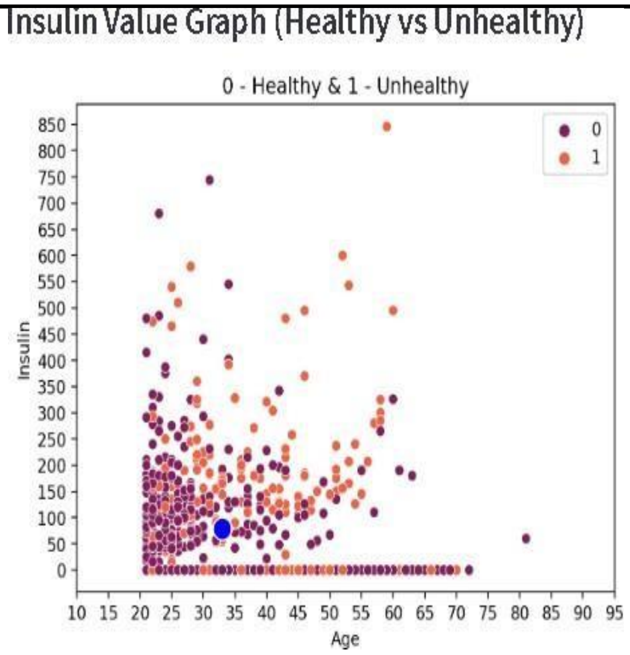

Figure 3.9 insulin value count

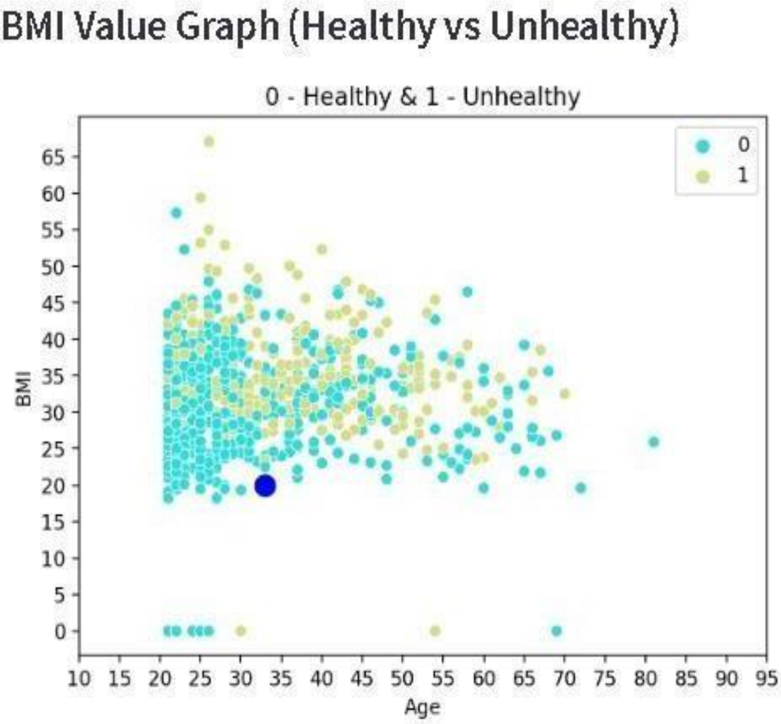

Figure 3.10 BMI Value Graph

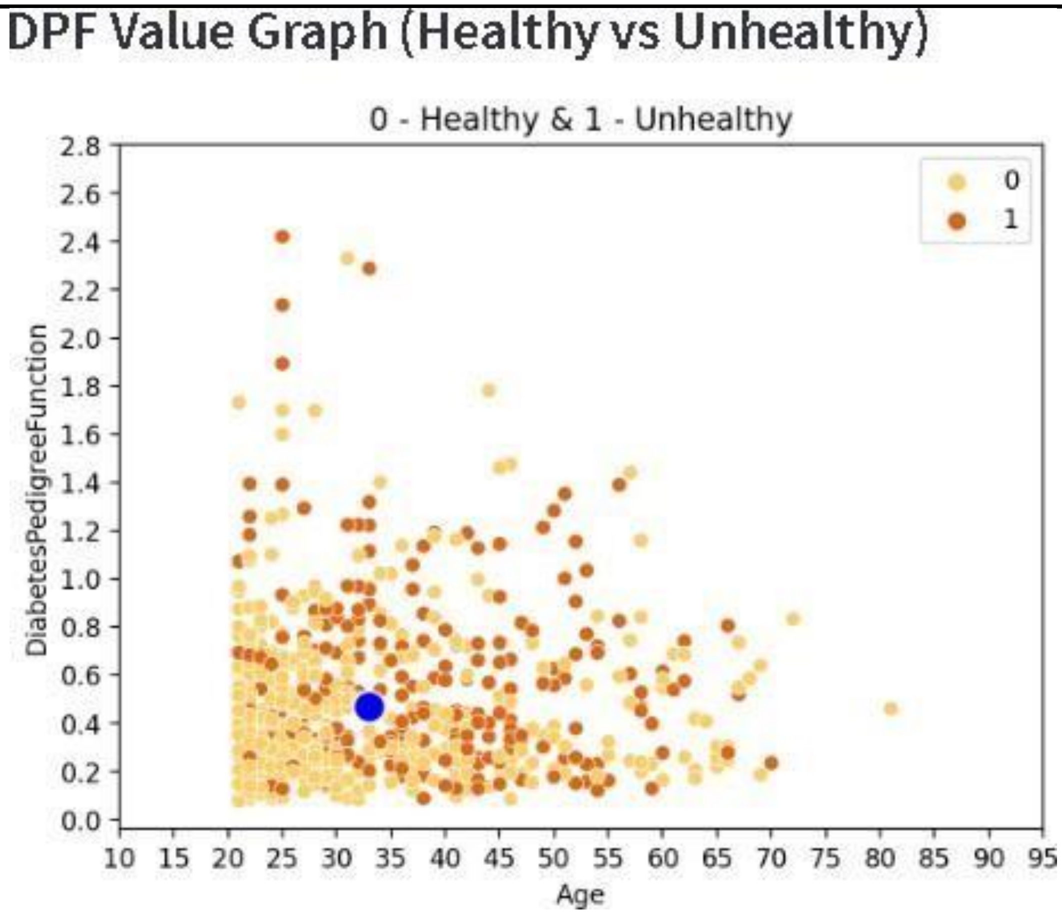

Figure 3.11 DPF value count

CHAPTER 4
CONCLUSION

CONCLUSION

Prediction of diabetes in pregnant women and neonatal mellitus in newborn children using machine learning revolves around developing predictive models to predict the risk of gestational diabetes in women expectant mothers and the likelihood of neonatal diabetes in newborns. By leveraging machine learning algorithms and analyzing various factors such as medical history, clinical parameters, lifestyle, and genetic predisposition, these models aim to provide early detection and personalized risk assessment. The goal is to enable healthcare professionals to intervene proactively, implement preventive measures. Various machine learning algorithms are applied to the dataset in this study, and classification is done using various algorithms, with Random Forest providing the highest accuracy. We've seen how machine learning algorithm accuracies compare to datasets. This research could be expanded to see how likely non-diabetic people are to develop diabetes in the coming years. As a result, we use common classification and ensemble algorithms on the dataset to make predictions.

Diabetes is a prevalent chronic disease that can be extremely dangerous to one's health. Diabetes is diagnosed when blood glucose levels are greater than normal, which is caused by insulin secretion or biological factors. Diabetes can harm our bodies in a variety of ways, including causing tissue, kidney, eye, and blood artery dysfunction. Based on physical examination data and consultation with doctors, machine learning may make a preliminary diagnosis of diabetes mellitus. Many techniques, including machine learning methods like Random Forest, Support Vector Machine, Decision Tree, and others, have recently been utilized to predict diabetes.

REFERENCE

REFERENCE

1 Марко Романович Басараб , Екатерина Олеговна Іванько , Вішвеш Кулкарні, "Prediction of the Development of Gestational Diabetes Mellitus in Pregnant Women Using Machine Learning Methods".(2021)DOI-10.20535/25234455.mea.228845

2 Zahura Zaman , Md. Ashrak Al Arif Shohas , Mahedi Hasan Bijoy , Meherab Hossain , Shakawat Al Sakib,"Assessing Machine Learning Methods for Predicting Diabetes among PregnantWomen".(2022)DOI-10.31632/ijalsr.2022.v05i01.005

3 AVSrinivas , Abbireddy Ramya , GT Chandralekha , Bhandaram Vaagdevi , K Anand

4 Quan Zou , Kaiyang Qu , Yamei Luo , Dehui Yin, Ying Ju , Jun Zhang, "Predicting Diabetes Mellitus With Machine Learning Techniques",(2018)DOI-10.3389/fgene.2018.00515

5 ,"Prediction of gestational diabetes based on explainable deeplearning andfogcomputing"(Lili Wei , Yueshuai Pan , Yan Zhang,"Application of machine learning algorithm for predicting gestational diabetes mellitus in early pregnancy"(2021),DOI10.2478/fon- 2021-0022

6 Thach Tran , Jane E.Hirst,"Prognosis Model for Gestational Diabetes Using Machine Learning Techniques"(2021),DOI10.2337/dc12-1418

7 Sara L. White , Debbie A Lawlor , Annette Briley,"Early Antenatal Prediction of Gestational Diabetes in Obese Women"(2016),DOI10.1371/journal.pone.0167846

8 Sumathi Amarnath , Meganathan Selvamani,"Prognosis Model for Gestational Diabetes Using Machine Learning Techniques" (2021),DOI10.18494/sam.2021.3119

9 M Sekulić , D Relić , "The role of machine learning in identification of early gestational diabetes mellitus prediction models"(2020),DOI10.26800/lv-142-supp5- 61

10 Xinxi Lu , Jikai Wang ," Prediction of Gestational Diabetes and Hypertension based on Pregnancy Examination Data"(2022), DOI10.1142/s0219519422400012

11 Jie Zhang , Fang Wang,"Prediction of Gestational Diabetes Mellitusunder Cascade and Ensemble learning Algorithm" (2022),DOI10.1155/2022/3212738

12 Nora El-Rashidy , Nesma E. ElSayed2022),DOI10.1007/s00500-022-07420

13 Yang Liu , Hua Sun,“Prediction Method of Gestational Diabetes Based on Electronic Medical RecordData”(2021)DOI10.1155/2021/6672072-
14 Jenny Yang , David Clifton,“Machine Learning-Based Risk Stratification for Gestational Diabetes Management”(2022),DOI10.3390/s22134805-
15 Hang Lai , Huaxiong Huang , Karim Keshavjee , Aziz Guergachi , Xin Gao “Predictive models for diabetes
mellitus using machine learningtechniques” (2019)DOI10.1186/s12902-019-0436-