International Journal of Advanced Scientific Innovation
ISSN: 2582-8436
Volume 06 Issue 10, September 2024

A Network Approach To Predict GDM Risk In Pregnant Women

Yashodha H R
M.Tech Student
yashodhahr1234@gmail.com
9880241470

Dr. Charan K V
Associate Professor
Dept of ISE, SIET, Tumkur
charan.kv@shrideviengineering.org
8147420383

Abstract

High blood sugar during pregnancy known as gestational diabetes mellitus (GDM) can cause difficulties for both the mother and the unborn child. Particularly in places where prenatal care is scarce, early detection and management are essential. This study suggests a combined machine learning prediction model to determine which expectant mothers are susceptible to gestational diabetes mellitus. We examined eight distinct models, incorporating deep learning methodologies.(Artificial Neural Networks) and conventional machine learning algorithms (Support Vector Machine, Naive Bayes, Random Forest, and Logistic Regression), using a dataset of 3526 pregnant women from Kaggle's Gestational Diabetes Mellitus dataset. With accuracy rates ranging from 87% to 97%, these models demonstrate the immense potential of machine learning to enhance GDM screening and management, especially in resource-constrained environments.

I. Introduction

When it comes to evaluating blood glucose (BG) control and metabolic health in diabetic patients, postprandial glycemic response (PPGR) is essential. The importance of keeping post-meal blood glucose levels within the normal range is emphasized by clinical research. Even with better metabolic control, diabetes pregnancies continue to pose a serious risk for problems with fetal development. Pregnancy deformities can be less common if ideal fasting glucose levels are maintained. Using a variety of machine learning techniques, recent research has concentrated largely on predicting blood glucose levels in individuals with type 1 diabetes. [1]

HIV/AIDS, cancer, diabetes, renal syndrome, inflammatory bowel disease, and cardiovascular disease are a some of the conditions linked to the complicated prevalence of anemia [8]. Hemoglobinopathies, malaria, and bilharzia are also important causes [8,9]. There are numerous forms of anemia, including iron deficiency, vitamin or iron deficiency, aplastic anemia, sickle cell disease, and thalassaemia. Every type of anemia has a variety of reasons, ranging from moderate to severe and temporary to permanent. Practical obstacles to

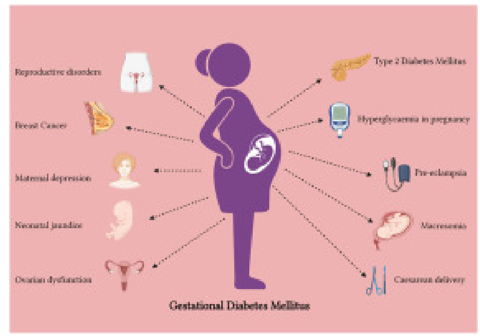

Fig. 1. Gestational diabetes

the laboratory process for identifying and detecting anemia in response to clinical concerns include a lack of technical expertise and inadequate financing for medical testing. [2]

The Global Diabetes Map (9th edition) by the International Diabetes Federation reports that the number of women experiencing diabetes is increasing worldwide during pregnancy, with approximately 20.4 million (15.8%) women affected by hyperglycemia, and 83.6% of these cases attributed diabetes mellitus during pregnancy (GDM). GDM, a common metabolic disease, typically resolves after delivery but poses significant risks during pregnancy. Pregnant women with GDM have a higher chance of face adverse outcomes that can complicate delivery.In order to screen for gestational diabetes mellitus (GDM), an oral glucose tolerance test (OGTT) that measures fasting plasma glucose levels one to two hours following the injection of glucose is typically advised between weeks 24 and 28 of pregnancy. To identify the risk of hyperglycemia early on, The American Diabetes Association advises against doing so. during the first prenatal visit. International organizations, however, disagree on a number of issues, including universal versus selective screening, when to test (early pregnancy versus 24-28 weeks), whether to use a one-step or two-step technique, and which diagnostic criteria to use.[3]

In this randomized trial evaluating the efficacy of treatment versus no treatment for mild gestational diabetes mellitus, researchers found that while rates of the primary outcome (a composite of stillbirth or perinatal death and neonatal complications) did not differ significantly between the two teams, significant reductions were observed in several secondary outcomes. This study sheds light on the potential benefits of treatment in managing specific complications associated with mild gestational diabetes mellitus. [4]

The World Health Organization (WHO) [12] and the UK National Institute for Health and Care Excellence (NICE) have validated the 75-gram oral glucose tolerance test (OGTT) as the preferred diagnostic test for gestational diabetes mellitus (GDM). As stated by the WHO/IADPSG criteria, GDM is identified if a woman exhibits a fasting plasma glucose level of 5.1 mmol/L or higher, a 1-hour postprandial glucose level of 10.0 mmol/L or above, or a 2-hour postprandial glucose level of 8.5 mmol/L or above. However, these diagnostic thresholds lack global consensus, resulting in varying practices in diagnosis and management across countries. Following diagnosis, the management of glycemic levels in GDM relies on self-monitoring of glucose levels, typically through fingerstick capillary blood tests. Different organizations advocate for distinct blood glucose targets in ladies who have GDM.M. [5]

Gestational diabetes mellitus (GDM) is a hyperglycemic disease that is first identified during pregnancy, characterized by blood glucose levels that are lower than those considered diagnostic for overt diabetes outside of pregnancy. Both the short- and long-term health of a mother and her kid are negatively impacted by GDM. For instance, women with GDM are more likely to develop type 2 diabetes or prediabetes over the long run as well as short-term preeclampsia. Similarly, children of women who develop prenatal diabetes mellitus are more susceptible to obesity or poor glucose tolerance later in life, as well as to macrosomia or hypoglycemia soon after delivery. Surprisingly, GDM has been linked to a great deal of additional negative outcomes. [6]

With a substantial concentration in South and South East Asia, diabetes mellitus during pregnancy (GDM) affects over 20 million live births worldwide, posing a serious and expanding worry. The diagnostic landscape of GDM has been historically contentious, marked by variations in screening protocols and diagnostic standards in various geographical areas. This variability has prompted the broader term ”hyperglycemia in pregnancy,” including pre-gestational diabetes, gestational diabetes, and diabetes detected during pregnancy in addition to GDM. The benefits of controlling hyperglycemia in pregnant women have been well-documented over the past ten years. These benefits extend to milder cases of hyperglycemia as well, as managing it can improve pregnancy outcomes and lower the risk of conditions like gestational hypertension and preeclampsia. Moreover, the ongoing association between low mother glucose levels and unfavorable neonatal outcomes emphasizes the significance of efficient management techniques. Targeted therapies for women with GDM and their families present a compelling area for concentrated healthcare efforts, despite the fact that GDM is frequently perceived as a temporary ailment. This is because it may have long-term effects on the cardiometabolic health of the mother and her kids. [7]

In recent years, the global prevalence of diabetes, particularly among women during pregnancy, has spurred significant research into the interplay between glycemia and pregnancy outcomes. The World Health Organization's (WHO) diagnostic criteria for hyperglycemia in pregnancy, established in 1999, were deemed in need of revision given the evolving landscape of evidence-based guidelines. Through systematic reviews and the use of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) methodology, efforts were made to reassess diagnostic cut-off values for gestational diabetes. This update not only addresses the diagnostic challenges of gestational diabetes but also sheds light on the elevated risks associated with hyperglycemia during pregnancy, including heightened chances of macrosomia, pre-eclampsia, and hypertensive disorders. Notably, interventions targeting gestational diabetes have demonstrated effectiveness in mitigating these risks, underscoring the critical role of accurate diagnosis and timely management in enhancing the health of both the mother and the fetus. [8]

Regarding the field of predicting gestational diabetes mellitus (GDM), prior research has aimed to identify a threshold value of fasting plasma glucose (FPG) in the first trimester, often through large-scale investigations. Setting diagnostic criteria at an FPG level of 6.1 mM or higher achieves nearly perfect specificity, but with an extremely low sensitivity, making it less feasible. Recent advancements have introduced novel biomarkers like angiopoietin-like protein 8, plasma fatty acid-binding protein 4, and various adipokines as potential predictors of GDM. However, their limited availability in clinical settings restricts their widespread use. Combining common risk factors such as advanced maternal age, body mass index (BMI), and using diabetic family history in prediction models presents a viable strategy. The accuracy of GDM prediction has significantly improved with the use of artificial intelligence, especially supervised machine learning (ML) approaches. However, there is a limited window of time for medical intervention because GDM forecasts usually materialize in the second trimester. [9]

In the last ten years, scholars have investigated using machine learning toestimate the risk of gestational diabetes (GDM) early. A recent meta-analysis study, published in December 2021, has identified key parameters for these models and analyzed various prognostic models for GDM risk prediction. The study included 25 trials involving women over 18 lacking a background in major diseases.

Machine learning models achieved an area under the receiver operating characteristic curve (AUROC) of 0.8492. The pooled sensitivity was 0.69 (95% CI 0.68-0.69; P.001; I^2=99.6%), and the pooled specificity was 0.75 (95% CI 0.75-0.75; P.001; I^2=100%). Logistic regression, one of the most commonly used ML approaches, had a pooled AUROC of 0.8151, while non-logistic regression models surpassed this with an AUROC of 0.8891. The four most commonly utilized variables in models created using different feature selection techniques were BMI, fasting blood glucose, maternal age, and family history of diabetes. Comparing machine learning approaches to conventional screening techniques, the study found that the former show promise in predicting GDM. To encourage their use, it is imperative to stress the necessity of reliable assessments and standardized diagnostic standards. [10]

The oral glucose tolerance test, or OGTT, has a high false positive rate, requires a lot of time from both patients and clinicians, and is difficult to apply to the whole population [14]. Pre-analytical laboratory techniques have the potential to greatly affect the outcomes. For example, glucose levels can be lowered by five to seven percent per hour by room temperature glycolysis by leukocytes and erythrocytes before centrifugation [15]. When the centrifugation process was used within ten minutes of sample collection in a recent Australian trial involving 12,317 women, the GDM diagnosis rate virtually quadrupled from 11.6% to 20.6% using the IADPSG criteria.

II. Literature Survey

In the past two decades, significant technological advancements have sparked in-depth investigation into leveraging artificial intelligence (AI), telemedicine, and mobile health to enhance healthcare delivery. These technologies play a vital part in managing chronic diseases, facilitating remote specialist care, and improving therapeutic outcomes under healthcare professionals' guidance, offering either fully automated or semi-automated support. Recent studies have particularly focused on utilizing mobile health tools and robots to treat long-term health issues including diabetes and high blood pressure. Xiong et al. used support vector machines (SVM) and light gradient boosting machines (lightGBM) to develop a risk identification technique for gestational diabetes mellitus (GDM) during the first 19 weeks of pregnancy. Similarly, Zheng et al. used biochemical markers and a machine learning (ML) model to develop a simple way to predict GDM in Chinese women during early pregnancies. Shen et al. investigated the possibilities of cutting-edge AI techniques for GDM assessment in situations with a shortage of physicians and clinical equipment, leading to the creation of an AI-powered app. [11]

Many machine learning approaches, such as dimensionality reduction and cross-validation methods such as ANN, AB, LR, DT, GPC, SVM, LDA, QDA, NB, and and RF, are being used in extensive research to predict diabetes. By locating and eliminating outliers, adding missing information, and attaining Researchers have enhanced the performance of machine learning (ML) models to an area under the curve (AUC) of up to 0.930. With an AUC of 0.819, Naive Bayes classifiers outperform decision trees and support vector machines (SVMs) in studies that examine their respective performances. In addition, studies have looked into meta-learning algorithms, the clinical use and explainability of various approaches, and the prediction of diabetes using algorithms such as CART, Adaboost, Logiboost, and graded learning. Furthermore, a condition that frequently appears in the second or third trimester of pregnancy is gestational diabetes (GDM).

In the previous ten years, researchers have investigated using machine learning to predict gestational diabetes mellitus (GDM) early. A recent meta-analysis study, released in December 2021, has identified key parameters for these models and analyzed various prognostic models for GDM risk prediction. The study included 25 trials involving women over 18 without a history of major diseases. Machine learning models achieved an area under the receiver operating characteristic curve (AUROC) of 0.8492. The pooled sensitivity was 0.69 (95% CI 0.68-0.69; P.001; I^2=99.6%), and the pooled specificity was 0.75 (95% CI 0.75-0.75; P.001; I^2=100%). Logistic regression,The pooled AUROC of one of the most widely used methods for machine learning was 0.8151, while non-logistic regression models outperformed it with an AUROC of 0.8891. The four most commonly utilized variables in models created utilizing various feature selection methods were BMI, fasting blood glucose, maternal age, and family history of diabetes. Comparing machine learning approaches to conventional screening techniques, the study found that the former show promise in predicting GDM. To encourage their use, it is imperative to stress the necessity of reliable assessments and standardized diagnostic standards. [12]

Women eligible for this study were those between 24 weeks 0 days and 30 weeks 6 days of gestation with ablood glucose levels that range from 135 to 200 milligrams per deciliter (7.5 to 11.1 mmol per liter) one hour post a 50-g glucose loading test. Exclusions were made for women with preexisting diabetes, abnormal glucose screening before 24 weeks of gestation, previous gestational diabetes, a history of stillbirth, multifetal gestation, asthma, or chronic hypertension; those taking corticosteroids; with known fetal anomalies; or with imminentor premature birth brought on by fetal or maternal problems. All participants gave written consent. consent, and the study was authorized by the relevant human subjects committees. Following an overnight fast, eligible women underwent a blinded 3-hour 100-g oral glucose-tolerance test. Samples were centrally analyzed, and the results were sent to the data coordinating center. Mild gestational diabetes mellitus was defined as fasting glucose levels below 95 mg

index	Age	No of Pregnancy	Previous Gestation	HDL
0	53.0	22	1	55.0
1	69	12.0	0	102.0
2	101.0	63	12.4	118.0
3	99.0	70	15.0	116.0

per deciliter (5.3 mmol per liter) and two or three timed glucose measurements exceeding set thresholds: 1-hour, 180 mg per deciliter (10.0 mmol per liter); 2-hour, 155 mg per deciliter (8.6 mmol per liter); and 3-hour, 140 mg per deciliter (7.8 mmol per liter). Random assignment to treatment or Control groups were carried out utilizing the simple urn method, stratified by clinical center. Treatment involved formal nutritional counseling, diet therapy, and insulin as needed, while the control group received usual prenatal care. Additionally, a cohort of women with a positive 50-g glucose loading test but normal oral glucose-tolerance test results was enrolled in the control group, matched by race and body-mass index. This inclusion allowed blinding regarding the diagnosis of mild pregnancy-related diabetes mellitus. Women with a fasting glucose level of 95 mg per deciliter or higher on the diagnostic oral glucose-tolerance test were excluded from the study. [13]

maternal characteristics like obesity, excessive gestational weight gain, and the presence of Gestational Diabetes Mellitus (GDM) significantly contribute to accelerated fetal growth and increased accumulation of adipose tissue. This often leads to the birth of infants classified asFor gestational age, large (LGA). Importantly, it should be noted that offspring can develop excess adiposity independently of birth weight. The consequences of being LGA in the context of GDM extend well beyond immediate concerns like birth injuries and neonatal hypoglycemia. Studies have highlighted the association between GDM-related LGA and subsequent childhood obesity, a troubling trend with global implications. Research indicates that by the age of two, around one in ten children are already obese, with projections suggesting that more than half will be obese by age 35. This persistence of obesity into adulthood significantly raisesthe likelihood of getting cardiovascular disease as well as type 2 diabetes These long-term implications underscore the complex interplay between maternal health during pregnancy, fetal development, and the future health outcomes of offspring, necessitating further research and proactive management strategies.

In the context of data mining and big data analytics, this section addresses a unique strategy for the identification of gestational diabetes by the use of machine learning methods. Pregnant women's information was gathered to forecast the likelihood of developing diabetes, which makes up the input data.

The data has undergone the following preprocessing steps:

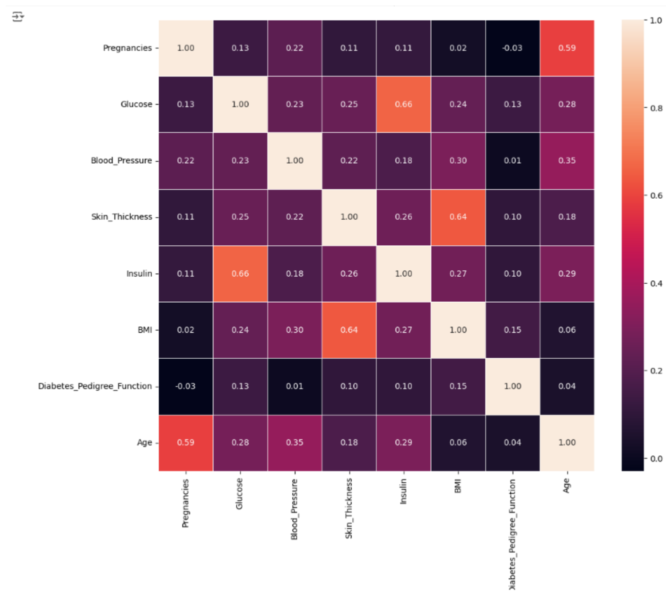

Fig. 2. graph

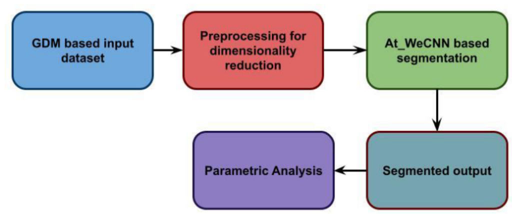

Fig. 3. flowchart

Dimensionality Reduction
The original dataset has been processed to reduce the number of features, making it more manageable for analysis.

Normalization
The data has been normalized to ensure that all features are on a similar scale, which is important for the subsequent machine learning algorithms.

Segmentation and Feature Fusion

Using a weighted convolutional neural network architecture based on an attention mechanism, the preprocessed input was divided into segments and the features were combined.

III. Methodology

This section describes a new machine learning approach for detecting gestational diabetes in the context of big data and data mining analytics. The input data, which was collected from pregnant women, was processed to normalize and minimize its dimensionality in order to predict diabetes.

The image provided combines a histogram with a probability density function, depicting the distribution of

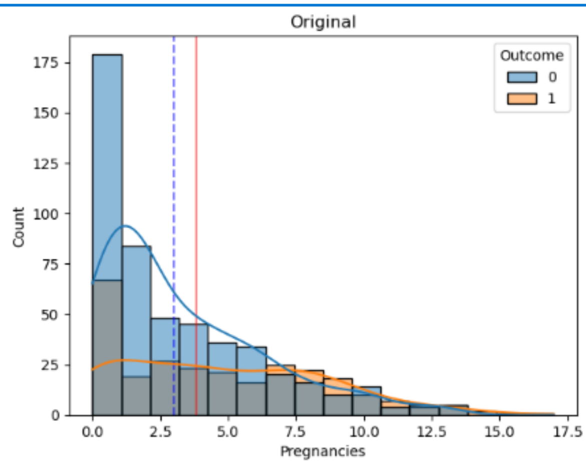

Fig. 4. flowchart

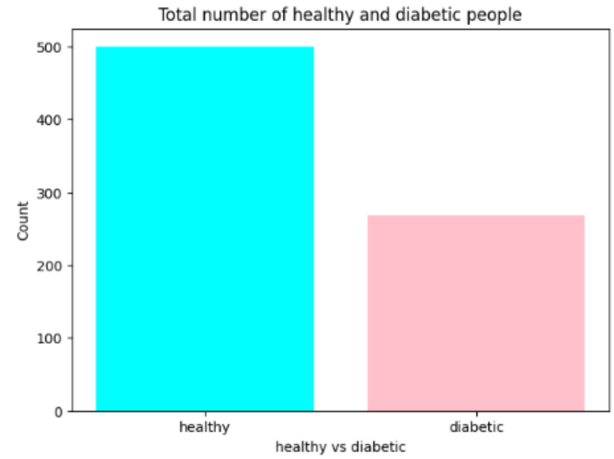

Fig. 5. healthy and diabetic chart

pregnancies within a group and distinguishing between outcomes denoted by blue (Outcome 0) and orange (Outcome 1), potentially indicating the absence or presence of gestational diabetes. This visualization reveals that most individuals in the group have had fewer than 5 pregnancies, with the count of those with Outcome 1 (gestational diabetes) consistently lower across all pregnancy counts. The distribution of pregnancies for both outcomes skews right, suggesting fewer individuals with higher pregnancy numbers. Transitioning toa machine learning strategy for gestational diabetes entails several steps: collecting a comprehensive dataset of potential features, preprocessing the data to handle missing values and standardize it, conducting exploratory data analysis to visualize distributions and explore relationships, selecting appropriate machine learning models for classification, training and evaluating these models, tuning hyperparameters for optimal performance, validating model generalization, interpreting results for feature importance, and potentially deploying the model in clinical settings for predictive and management purposes. However, it's crucial to note that the provided histogram alone lacks the depth necessary for predicting gestational diabetes through machine learning; a more extensive dataset with multiple features and a larger sample size would be required for a robust predictive model.

The provided image is a bar chart that depicts the total count of healthy and diabetic individuals in a specific dataset. The chart is bifurcated separated into two groups: "healthy" and "diabetic." The "healthy" category significantly surpasses the "diabetic" category, with a total of 500 individuals as opposed to approximately 268 people having a diabetes diagnosis. A note of observation beneath the chart underscores the imbalance in the dataset, with a larger number of healthy individuals compared to those with diabetes. This imbalance could potentially skew any statistical analysis or machine learning model training conducted using this dataset. Techniques such as resampling may be necessary to ensure that the analysis or predictive modeling is not biased towards the majority class, which in this case, are the healthy individuals.

The provided heatmap depicts the correlation among various health-related variables likely used in a study of gestational diabetes employing machine learning. These variables include Glucose, Blood Pressure, Skin Thickness, Insulin, and BMI (Body Mass Index). The heatmap utilizes a color scale where blue signifies a positive correlation and red indicates a negative correlation, with intensity reflecting correlation strength. Notably, the strongest correlation is between Insulin and Skin Thickness, showing a moderately strong positive correlation of 0.7, suggesting a tendency for Insulin levels to increase with higher Skin Thickness, or vice versa. Additionally, there are weaker positive correlations of 0.3 between BMI and Skin Thickness, and 0.1 between BMI and Insulin, implying some association between higher BMI, greater Skin Thickness, and higher Insulin levels. However, other correlations displayed are close to zero, indicating little linear relationship. An observation in the heatmap notes a clear relationship between Skin Thickness and Insulin, supported by their highest correlation value. It's important to note that correlation doesn't imply causation, necessitating further analysis to comprehend these relationships and potential influencing factors.

IV. CONCLUSION

In conclusion, the realms of AI and ML present promising and emerging avenues for monitoring and managing gestational diabetes among women. While ML and AI have demonstrated their utility in studies and medical settings, aiding in patient monitoring through risk stratification, discovering patient subgroups, and predicting outcomes using natural language processing [119–130], similar strategies tailored for GDM remain underdeveloped. The dynamic arena

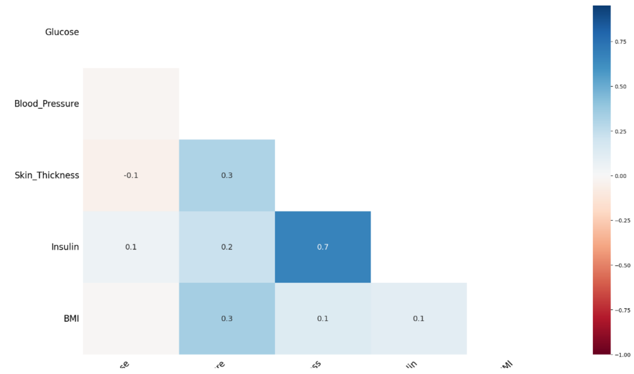

Fig. 6. Predictive Analysis of Gestational Diabetes

of data collection from diverse sensors, encompassing activity tracking, quantifying food intake, blood glucose monitoring, and medication management, holds potential for advancing GDM management. Yet, there persist numerous unresolved inquiries for data scientists, engineers, and clinicians. The imperative for personalized, explainable, and trustworthy AI and ML models is paramount, aiming to support patients and clinicians in enhancing lifestyles and achieving favorable short-term and long-term clinical results. It is critical to swiftly develop digital health technologies and elucidate AI methodologies to identify patients in varying risk categories early on (preventive medicine) and furnish clinicians with predictive monitoring models for devising reactive treatment plans.

Deep learning models and other machine learning-based models are presented in this study paper for the prediction and classification of gestational diabetes mellitus (GDM). Four stages of preparation were applied to the medical data: format conversion, class labeling, normalization, and replacement of missing values. The machine learning model was then trained using the preprocessed data to identify the correct class label. SVM, Naive Bayes, Random Forest, Logistic Regression, XGBOOST, Decision Tree, SGD, and ANN are the eight models that were created for early prediction. The suggested model outperformed the most recent research study findings in terms of accuracy.

References

[1] e elena a and anopova, anna d and kokina, maria a and gorelova, inga v and pervunina, tatiana m and grineva, elena n and popova, polina v. *Ieee Access*, 8:219308–219321, 2020.

[2] MOHAMMED ALOTAIBI and NOJOOD ALJEHANE. Early prediction of gestational diabetes using machine learning techniques. *Journal of Theoretical and Applied Information Technology*, 101(21), 2023.

[3] K. Arun Kumar, R. Rajalakshmi, Shashikala H. K., Maithli Ganjoo, Aman Vats, and Rajneesh Tyagi. Machine learning algorithm for the detection of gestational diabetes: Research challenges with big data and data mining. *International Journal of Intelligent Systems and Applications in Engineering*, 10(2s):260 –, Dec. 2022.

[4] M.B. Landon et al. A multicenter, randomized trial of treatment for mild gestational diabetes. *New England Journal of Medicine*, 361(14):1339–1348, October 2009.

[5] Huiqi Y Lu, Xiaorong Ding, Jane E Hirst, Yang Yang, Jenny Yang, Lucy Mackillop, and David A Clifton. Technologies in digital health and machine learning for managing gestational diabetes and blood glucose monitoring. *IEEE reviews in biomedical engineering*, 17:98–117, 2023.

[6] Daniela Mennickent, Andrés Rodríguez, Marcelo Farías-Jofré, Juan Araya, and Enrique Guzmán-Gutiérrez. A review of machine learning-based algorithms for predicting gestational diabetes mellitus before 24-28 weeks of pregnancy. *Artificial Intelligence in Medicine*, 132:102378, 2022.

[7] P. Saravanan et al. Gestational diabetes: opportunities for improving maternal and child health. *Lancet Diabetes Endocrinology*, 8(9):793–800, September 2020.

[8] World Health Organization. Who guidelines for the diagnosis and categorization of hyperglycemia initially identified during pregnancy. *Diabetes Research and Clinical Practice*, 103(3):341–363, March 2014.

[9] Yan-Ting Wu, Chen-Jie Zhang, Ben Willem Mol, Andrew Kawai, Cheng Li, Lei Chen, Yu Wang, Jian-Zhong Sheng, Jian-Xia Fan, Yi Shi, et al. Using sophisticated machine learning, early detection of gestational diabetes mellitus in the chinese population. *The Journal of Clinical Endocrinology & Metabolism*, 106(3):e1191–e1205, 2021.

[10] Yan Xiong, Lu Lin, Yu Chen, Stephen Salerno, Yi Li, Xiaoxi Zeng, and Huafeng Li. Predicting gestational diabetes mellitus in the first 19 weeks of pregnancy using machine learning algorithms. *The journal of pediatric and maternal-fetal medicine*, 35(13):2457–2463, 2022.

[11] Zheqing Zhang, Luqian Yang, Wentao Han, Yaoyu Wu, Linhui Zhang, Chun Gao, Kui Jiang, Yun Liu, and Huiqun Wu. Machine learning prediction models for gestational diabetes mellitus: a meta-analysis. *Journal of medical Internet research*, 24(3):e26634, 2022.

[12] Zheqing Zhang, Luqian Yang, Wentao Han, Yaoyu Wu, Linhui Zhang, Chun Gao, Kui Jiang, Yun Liu, and Huiqun Wu. Machine learning prediction models for gestational diabetes mellitus: Meta-analysis. *Journal of Medical Internet Research*, 24(3):e26634, 2022.

[13] W.W. Zhu et al. Beijing's high incidence of gestational diabetes mellitus: impact of mother birth weight and additional risk factors. *Chinese Medical Journal*, 130(9):1019–1025, May 2017.