International Journal of Basic Science and Technology
December, Volume 10, Number 3, Pages 409 - 417 http://www.ijbst.fuotuoke.edu.ng /
https://doi.org/ 10.5281/zenodo.15123405 s ISSN 2488-8648

Artificial Intelligence Assisted Gestational Diabetes Mellitus Monitoring System: An Experimental Approach

Ogunmodimu, Dupe Catherine
Department of Computer Science and Informatics, Federal University Otuoke
Bayelsa State, Nigeria.

Article Information

Article # 10053
Received: 6th Nov. 2024
Revision:4th Dec. 2024
2nd Revision: 11th Dec. 2024
Acceptance 18th Dec. 2024
Available online:
27th December, 2024.

Key Words
Clinical Decision Support System,
Artificial Intelligence,
Gestational Diabetes Mellitus,
Machine Learning

Abstract

Gestational diabetes Mellitus, often referred to as pregnancy-related diabetes, is a type of diabetes that occurs during pregnancy. It can develop at any stage and is characterized by high blood sugar levels. If left undetected and untreated, gestational diabetes mellitus can result in complications for both the mother and the unborn child before, during, and after birth. The primary objective of this study is to develop and evaluate predictive models that can identify pregnant women at risk for gestational diabetes, this will help identify high-risk mothers who require earlier treatment, monitoring, and medication. The clinical decision support system proposed in this study for diagnosing Gestational Diabetes Mellitus was developed using the electronic data of routine antenatal care of pregnant women obtained from the UCI machine learning data repository. We employed various machine learning algorithms, including Logistic Regression, K-Nearest Neighbors, Support Vector Machine, Random Forest Classifier, and Gradient Boosting Classifier, to develop and evaluate classification models aimed at early detection of gestational diabetes in pregnant women. The models were trained using 70% of the data and validated with the remaining 30%. To assess the performance of these predictive models, we compared them based on several evaluation metrics, including accuracy, recall, precision, and the AUC score. In the validation dataset, the Support Vector Machine (SVM) model outperformed other classifiers, achieving an accuracy of 98%, a recall of 95%, a precision of 100%, and an AUC score of 100%. An exploratory analysis of the Gestational Diabetes Mellitus (GDM) dataset identified several factors associated with an increased risk of developing Gestational Diabetes Mellitus, including age, the number of pregnancies, diastolic blood pressure, gestation in previous pregnancies, and family history. The results from the Support Vector Machine model demonstrated high accuracy, interpretability, and superiority in predicting Gestational Diabetes Mellitus using the GDM dataset.

***Corresponding Author: Ogunmodimu, D.C**.; dupeogunmodimu@gmail.com

Introduction

Gestational diabetes mellitus is one of the most common medical complications associated with pregnancy, and its incidence is steadily increasing worldwide. The World Health Organization defines gestational diabetes mellitus as “glucose intolerance or hyperglycemia that is first recognized or appears during pregnancy.” This condition arises when blood sugar levels become excessively high (hyperglycemia) due to hormonal changes that occur during pregnancy. During pregnancy, the placenta produces hormones such as human placental lactogen, cortisol, and estrogen. These hormones can hinder the action of insulin, which is essential for regulating glucose levels in the blood, insulin helps the body utilize glucose for energy and keeps blood sugar levels within a normal range. When the body cannot produce enough insulin to offset the increased resistance caused by these hormones, it leads to elevated glucose levels in the bloodstream.

Undetected or poorly managed gestational diabetes mellitus can pose serious risks for both the mother and the developing fetus. Potential complications for the mother include a heightened risk of preeclampsia, the need for cesarean delivery, and the likelihood of developing type 2 diabetes later in life. For the baby, gestational diabetes mellitus can result in excessive weight gain (macrosomia), preterm birth, and hypoglycemia after birth, among other issues.

Monitoring and managing blood sugar levels through dietary modifications, regular physical activity, and, in some cases, insulin therapy, are crucial for ensuring a healthy pregnancy outcome. Early detection and

International Journal of Basic Science and Technology
December, Volume 10, Number 3, Pages 409 - 417 http://www.ijbst.fuotuoke.edu.ng /
https://doi.org/ 10.5281/zenodo.14510094 **ISSN 2488-8648**

treatment of gestational diabetes are essential to minimizing health risks for both mother and child. (Cleveland Clinic, 2024).

In the past two decades, the prevalence of gestational diabetes has significantly increased, largely due to lifestyle changes and a rise in maternal age. Studies suggest that the rates of gestational diabetes range from 1% to 26%. In Asia, the incidence varies between 3% and 21.2%, while globally it ranges from 0.31% to 18% (Frontino, 2023). Some rural areas in Africa have reported a prevalence as low as 3.7%. In contrast, a crude prevalence of 13.9% was found among women at high risk for gestational diabetes in an urban population in Nigeria (Mba et al.2023).

Mothers diagnosed with gestational diabetes mellitus face a heightened risk of several serious complications during pregnancy and afterward. These complications include preeclampsia, which is characterized by high blood pressure and potential damage to other organ systems; placental dysfunction, which can impact the flow of nutrients and oxygen to the fetus; and metabolic disturbances that may affect both maternal and fetal health. Additionally, women with gestational diabetes mellitus are more likely to require a cesarean delivery due to complications related to the condition. Furthermore, those who experience gestational diabetes mellitus have a slightly increased risk of developing Type 2 diabetes later in life, especially if they do not lose the excess weight gained during pregnancy or have a family history of diabetes.

For the baby, gestational diabetes can lead to several dangerous outcomes. These include increased birth weight, known as fetal macrosomia, which can complicate delivery and heighten the risk of birth injuries. Newborns may also face respiratory issues at birth due to delayed lung development associated with maternal diabetes. Additionally, hypoglycemia—an abnormally low blood sugar level—can occur shortly after birth, potentially resulting in seizures and long-term neurological issues. Long-term risks for the child also include obesity, which can begin in childhood, premature birth, and a significantly increased likelihood of developing Type 2 diabetes as they grow older.

Overall, gestational diabetes presents significant risks, emphasizing the need for early diagnosis, diligent monitoring, and effective management throughout pregnancy to achieve the best outcomes for both mother and child.

Artificial Intelligence (AI) is increasing, revolutionizing, and streamlining data processing and decision-making in healthcare. A subset of AI, machine learning, can create reliable systems for healthcare diagnostics and forecasts. With the advancements in machine learning, it has become feasible to detect gestational diabetes and enhance the understanding of the condition.

The primary objective of this study is to leverage machine learning models to predict the occurrence of gestational diabetes and to rigorously assess their performance. Early detection of gestational diabetes is vital as it plays a significant role in managing and mitigating potential complications for both the mother and the infant, including risks such as preeclampsia, premature delivery, and long-term diabetes susceptibility.

With the advancement of computing technologies and the growing availability of well-curated and labeled datasets in the healthcare sector, machine learning has become an indispensable tool for enhancing diagnostic capabilities related to diabetes in pregnant women.

This study employed different machine learning algorithms, including Logistic Regression, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Random Forest Classifier, and Gradient Boosting Classifier. The selection of these algorithms is based on their ability to handle various data characteristics and their potential to yield valuable insights into the prediction of gestational diabetes.

Several data preprocessing techniques were implemented to optimize model performance. Mean imputation is utilized to address missing data points, thereby preserving the integrity of the dataset and minimizing potential bias. Furthermore, min-max normalization is applied for feature scaling, standardizing input features to a common scale and thereby enhancing both training efficiency and predictive accuracy. Through this comprehensive methodology, this study aims to provide significant contributions to the effective prediction of gestational diabetes, ultimately fostering improved healthcare decisions for pregnant women.

In recent decades, extensive research has firmly established the significance of early prediction of gestational diabetes, demonstrating its crucial impact on enhancing both maternal and fetal health outcomes, with a variety of methodologies employed to enhance diagnostic accuracy and patient outcomes, among these approaches, machine learning techniques have emerged as flexible and adaptable prediction algorithms. These algorithms leverage large datasets to identify complex patterns and relationships among variables, offering significant potential advantages over traditional statistical methods enhancing the accuracy and timeliness of gestational diabetes predictions, machine learning can promote early

International Journal of Basic Science and Technology
December, Volume 10, Number 3, Pages 409 - 417 http://www.ijbst.fuotuoke.edu.ng /
https://doi.org/ 10.5281/zenodo.14510094 **ISSN 2488-8648**

interventions that may result in improved health outcomes for both mothers and infants.

Artificial intelligence (AI) technology has recently been revealed to exhibit a strong capacity for self-learning with increased gestational diabetes mellitus prediction accuracy. This is especially true of supervised machine learning (ML) techniques.

Wu *et al.* (2021) created a deep neural network model to predict the risk of gestational diabetes mellitus (GDM) in Chinese women during the first trimester of pregnancy. They utilized an electronic medical record system to extract pregnancy data for 73 different variables collected during this period. Through a feature selection process enhanced by machine learning techniques, they identified 17 factors that could predict early GDM. From this initial set, they selected seven key variables based on their practical relevance to clinical settings. Using both the seven-variable model and the larger dataset of 73 variables, they applied advanced machine-learning methods to create models capable of predicting early GDM across various scenarios. The deep neural network model demonstrated impressive performance, achieving a significant discriminative capability with an area under the curve (AUC) value of 0.80.

Xiong et al. (2020) conducted an insightful implementation of eight widely used machine learning algorithms: Gradient Boosting Decision Tree (GBDT), AdaBoost, Voting Classifier, Logistic Regression, Gradient Boosting, LightGBM (LGB), Extreme Gradient Boosting (XGB), and Random Forest. They also applied two effective regression techniques: logistic regression and stepwise logistic regression with restricted cubic splines (RCS). The performance of these models was thoroughly evaluated using a variety of metrics. The findings revealed that the machine learning models and logistic regression achieved moderate performance on the validation dataset, with area under the curve (AUC) values ranging from 0.59 to 0.74. This suggests that while the models showed promise, there is room for improvement in their predictive capabilities.

Zhang *et al.* (2022) conducted a comprehensive evaluation of several machine learning algorithms, including XGBoost, LightGBM, Gradient Boosting, Decision Tree, and Logistic Regression. Their findings demonstrate that the models implemented achieved high levels of accuracy, recall, and precision. This highlights the effectiveness of these algorithms in enhancing the prediction and risk assessment of gestational diabetes mellitus (GDM).

In the study conducted by Rani (2020), a novel predictive system was developed for the early detection of diabetes. This system integrated various machine learning techniques and algorithms, including Logistic Regression, Random Forest, K-Nearest Neighbors, Support Vector Machine, and Decision Tree. The performance of each algorithm was evaluated based on its predictive accuracy. The findings indicated that the Decision Tree model outperformed the other algorithms assessed in predicting diabetes.

Jader and Aminifar (2022) introduced an integrated predictive framework to diagnose gestational diabetes. This model employs the K-Means clustering technique to facilitate data reduction and utilizes the elbow method for determining the optimal k-value. For the prediction phase, various classification methods are implemented, including, k-nearest neighbors (KNN), random forest, decision trees, support vector machine (SVM), logistic regression, and Naive Bayes. The authors reported that the merger of K-Means clustering, the elbow method, Mahalanobis distance, and an ensemble technique yields a notable improvement in prediction accuracy.

To ascertain whether a person has diabetes, Trivedi *et al*. (2022) used artificial neural networks. Another criterion employed in the training process was a neural network error function. Following neural network training, their findings showed that their suggested model had an average error of 0.01 and was 99.6% accurate in determining whether a person had diabetes.

Hou *et al*. (2020) constructed a LightGBM prediction model and compared its performance with that of Random Forest and XGBoost using the ROC curve. Their findings demonstrated that the AUC (Area Under the Curve) for the LightGBM model was 85.2%. These results indicate that the LightGBM model offers significant advantages and achieves better classification performance in predicting gestational diabetes mellitus compared to the other models.

Srivastava *et al*. (2019) estimated the occurrence of gestational diabetes using Microsoft Azure AI services. Their proposed classification model predicts the likelihood of gestational diabetes based on various factors identified in the early stages of pregnancy. To build and validate this predictive model, they utilized the Pima Indian dataset from the UCI Machine Learning Repository. The algorithm was tested on a dataset comprising 768 samples and achieved an accuracy of 77.8%.

Materials and Methods

The study considered four stages for developing a clinical decision support model for gestational diabetes diagnosis. The stages include data

International Journal of Basic Science and Technology
December, Volume 10, Number 3, Pages 409 - 417 http://www.ijbst.fuotuoke.edu.ng /
https://doi.org/ 10.5281/zenodo.14510094 ISSN 2488-8648

acquisition, preprocessing, model development, and evaluation of the developed model's performance on the test set. We employed feature scaling in this study due to its significance in improving the predictive model's performance.

Dataset Acquisition

The dataset used in this study is the Gestational Diabetes Mellitus (GDM Data Set), it is electronic data of routine antenatal care of pregnant women obtained from the UCI machine learning data repository, the dataset consists of 3525 instances with 17 features such as age, number of pregnancies, gestational in previous pregnancy, HDL, family history of diabetes, Sys BP, Dia BP, PCOS, large child or birth default, prediabetes, etc. Table 1 shows the dataset features and their respective data types with GBM and Non GDM as the target variables. The Dataset consists of 1,372 cases of GMB and 2,153 cases of Non GMB as shown in Figure 1

Feature	Data Type
Case Number	integer
Age	integer
No of Pregnancy	integer
Gestation in previous Pregnancy	integer
BMI	float
HDL	float
Family History	integer
unexplained prenetal loss	integer
Large Child or Birth Default	integer
PCOS	integer
Sys BP	float
Dia BP	integer
OGTT	float
Hemoglobin	float
Sedentary Lifestyle	integer
Prediabetes	integer
Class Label(GDM /Non GDM)	Integer

Table 1: Gestational Diabetes Mellitus (GDM Data Set). Features with Data Type

Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) was conducted on the dataset to uncover underlying patterns and trends. This process involved visualizing the data through various graphs and charts, as well as performing statistical analyses to summarize key characteristics. By examining distributions, correlations, and potential outliers, we aimed to gain a deeper understanding of the dataset. Figure 1 illustrates how the target variables are distributed in the GDM dataset.

International Journal of Basic Science and Technology
December, Volume 10, Number 3, Pages 409 - 417 http://www.ijbst.fuotuoke.edu.ng /
https://doi.org/ 10.5281/zenodo.14510094 ISSN 2488-8648

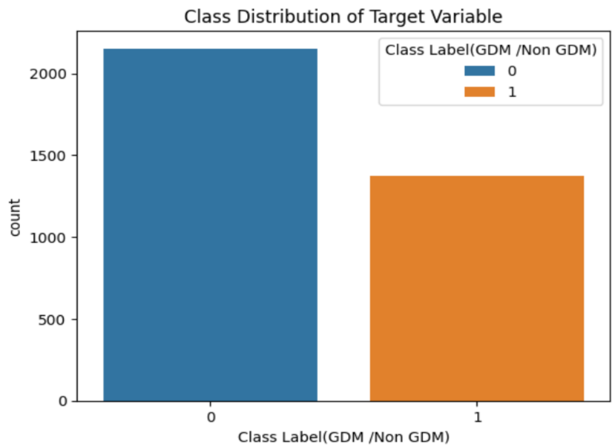

Figure 1: Class Distribution of the Target Variable

Figures 2, 3, and 4 graphically represent the distribution of the outcome by age group, Dia BP, and number of pregnancies. Figure 2 indicates that pregnant women aged 35 to 44 are at a higher risk of developing gestational diabetes. Figure 3 highlights that blood pressure is a significant factor influencing the development of gestational diabetes. Additionally, Figure 4 illustrates the distribution of outcomes based on the number of pregnancies; specifically, the likelihood of not developing gestational diabetes decreases as the number of pregnancies increases.

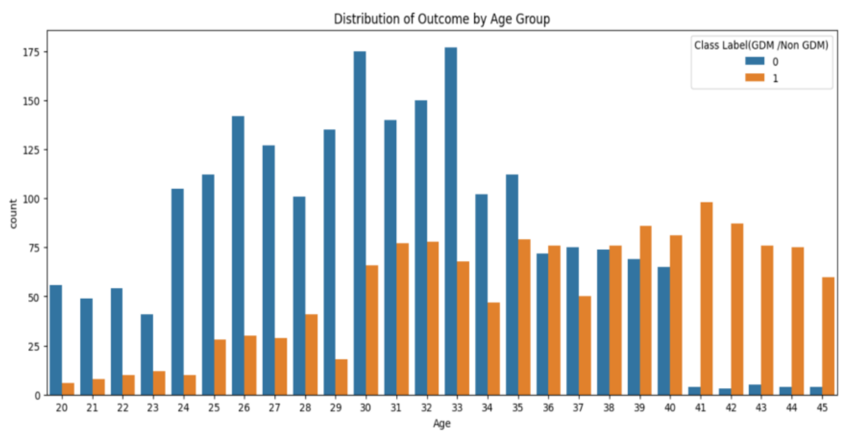

Figure 2: Distribution of Outcome by Age Group

International Journal of Basic Science and Technology
December, Volume 10, Number 3, Pages 409 - 417 http://www.ijbst.fuotuoke.edu.ng /
https://doi.org/ 10.5281/zenodo.14510094 **ISSN 2488-8648**

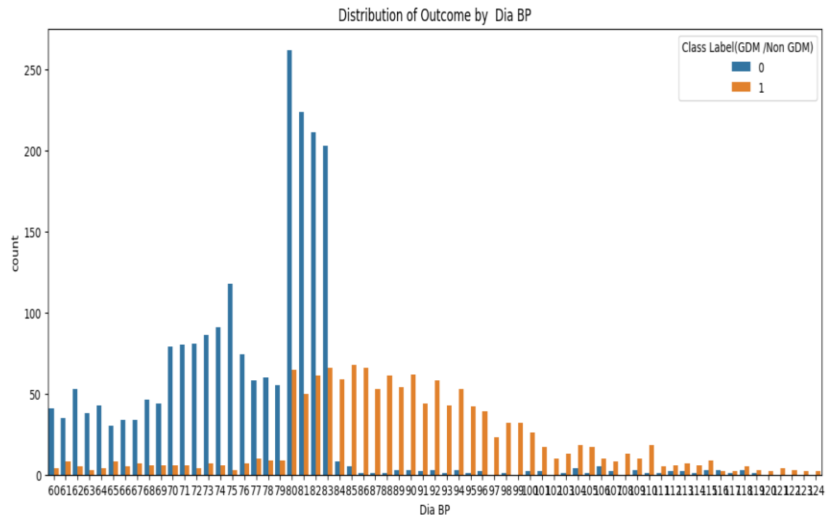

Figure 3: Distribution of Outcome by Dia BP

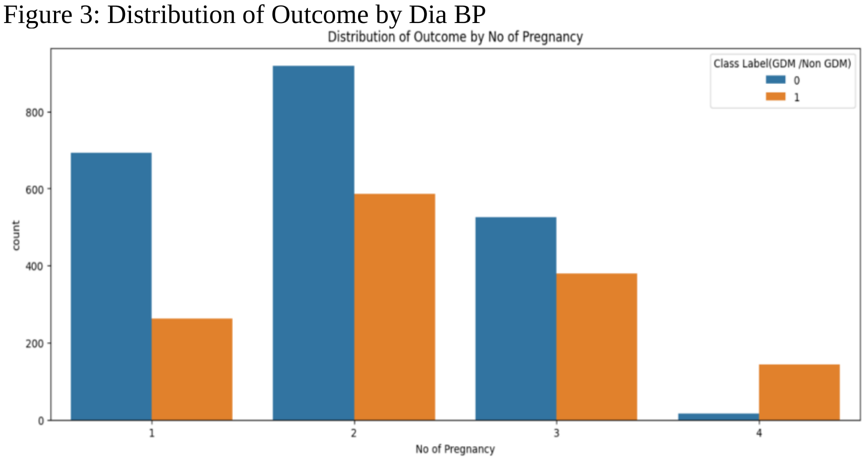

Figure 4: Distribution of Outcome by No of Pregnancy

The overall framework of the gestational diabetes mellitus risk prediction model is shown in Figure 5 below:

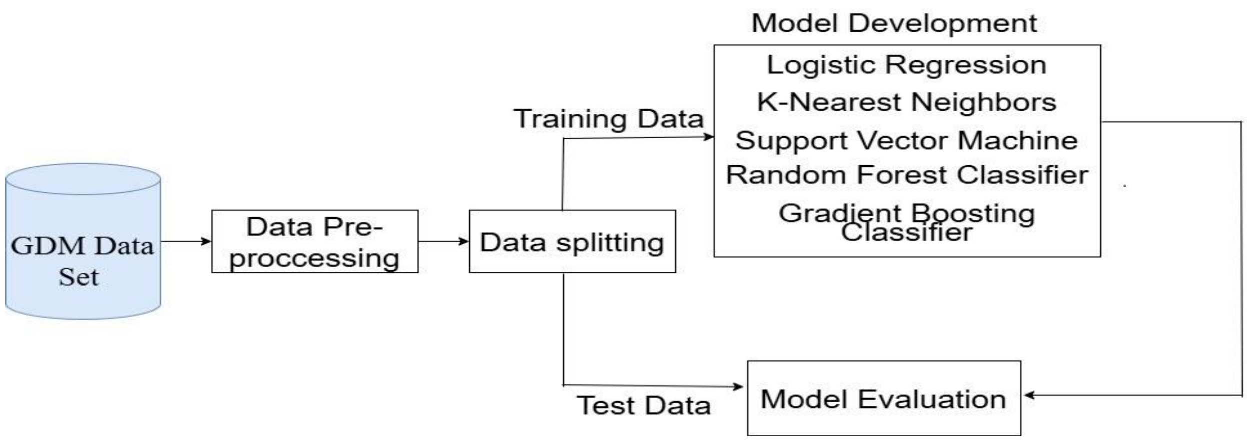

Figure 5: Gestational Diabetes Mellitus Risk Prediction Framework

Data Preprocessing

December, Volume 10, Number 3, Pages 409 - 417 http://www.ijbst.fuotuoke.edu.ng /
https://doi.org/ 10.5281/zenodo.14510094 **ISSN 2488-8648**

Data cleaning and transformation are important steps in developing effective machine-learning models. The dataset used in this study has a few missing values, encoded as blanks and NaNs. These missing values were handled before the actual processing. This study employed different techniques, such as dropping the entire column, and a mean imputation method which replaces the missing values with the mean of each column, to handle missing data values to increase the classification accuracy.
The mean imputation method is employed because repeatedly removing data from the dataset is impractical, doing so may significantly reduce the dataset's size, raising worries about biasing it and producing inaccurate results. We have dropped the Case Number column because it is irrelevant to model prediction.

Mean Imputation for Handling Missing Values

Handling missing data is an important step in the data preprocessing stage since it can have a major impact on the proposed model's correctness and reliability. Filling in missing values using the mean is a straightforward approach to estimating unknown values from existing data.
Imputation is a technique for replacing missing data with a substitute value to keep the majority of the data/information in the dataset. We used mean imputation to manage missing and inconsistent values in the dataset. This function replaces an attribute's value with its group's mean. The mean can be used to approximate certain properties. The following steps were followed to fill in the missing numbers with the mean:
Step 1: Identify columns or variables that have missing values.
Step 2: Calculate the mean for each column with missing values.
Step 3: Replace the missing values in each column with their respective mean values.

Normalization

Normalization is a scaling technique used in machine learning to modify the numeric column values of a dataset to a common scale. Normalization aims to transform features to be on a similar scale. In this study, we used Min-max normalization to normalize data using feature values and transformations. This method guarantees the same scale for all the features.
In Min-max normalization, for every feature, the minimum value is changed into 0, the maximum value is converted into 1, and every other value is converted into a decimal between 0 and 1, using equation 1.

$$Xn = (X - Xminimum) / (Xmaximum - Xminimum) \quad \text{equation 1}$$

Xn = Value of Normalization
Xmaximum = Maximum value of a feature
Xminimum = Minimum value of a feature
Normalization provides the following benefits:
1. It facilitates faster convergence of models during training. When features have varying ranges, gradient descent can "bounce," leading to slower convergence.
2. It helps models make more accurate predictions. However, the resulting model might produce slightly less useful predictions when different features have varied ranges.
3. To avoid the "NaN trap," it is important to understand that NaN, which stands for "Not a Number," arises when feature values exceed the limits of floating-point precision. When a value in a model is converted to NaN, it can lead to a cascading effect, resulting in additional values also becoming NaN. This phenomenon can significantly compromise the integrity and reliability of the model. Proper management of this issue is essential to maintain optimal model performance.
4. It allows the model to learn the appropriate weights for each feature. Without feature scaling, the model becomes biased towards attributes with large ranges while neglecting features with smaller ranges.
4. Model Development
Five machine learning algorithms were developed to compare the performance of different classification models in predicting gestational diabetes mellitus. The algorithms used include Logistic Regression, K-Nearest Neighbors, Support Vector Machine, Random Forest Classifier, and Gradient Boosting Classifier.
The dataset was divided into a training set comprising 70% of the data and a test set comprising 30%. The training set included 2,820 records of pregnant women, each with 15 features, while the test set contained 705 records, also with 15 features.

Model Evaluation

This study evaluated the overall performance of classification models using accuracy, recall, precision, and the area under the ROC curve (AUC) scores.

Results and Discussion

This section provides a detailed analysis of the performance of each model. The machine learning algorithms and visualization of results were implemented in Google Colaboratory using Python 3.10. We utilized accuracy, recall, precision, and the area under the Receiver Operating Characteristic (ROC) curve to compare the overall performance of

International Journal of Basic Science and Technology
December, Volume 10, Number 3, Pages 409 - 417 http://www.ijbst.fuotuoke.edu.ng /
https://doi.org/ 10.5281/zenodo.14510094 ISSN 2488-8648

the classification models. Accuracy serves as a generalized statistic for evaluating a classifier's performance. Precision indicates how effectively the classifier predicts positive outcomes, specifically measuring the classifier's ability to identify actual cases of gestational diabetes mellitus among all its predictions. A model with 100% precision would indicate that there are no false positives. Recall, on the other hand, is the percentage of true positive predictions among all positive forecasts. Ideally, a model with no false negative cases should achieve 100% recall. Receiver operating characteristic (ROC) curves visually represent classifier performance by plotting the true positive rate against the false positive rate. The area under the ROC curve (AUC) serves as a metric for assessing the effectiveness of the machine learning algorithms.

The modeling results for Logistic Regression, K-Nearest Neighbors, Support Vector Machine, Random Forest Classifier, and Gradient Boosting Classifier are presented in Table 2.

Table 2: Results for all Algorithms

Algorithms	Accuracy	Recall	Precision	AUC Score
Logistic Regression	97	96	95	98
K-Nearest Neighbors	97	97	96	99
Support Vector Machine	97	98	95	100
Random Forest Classifier	97	97	94	100
Gradient Boosting Classifier	97	96	95	100

The results of the experiment, as shown in Table 2, indicate that among the algorithms tested—Logistic Regression, K-Nearest Neighbors, Support Vector Machine, Random Forest Classifier, and Gradient Boosting Classifier—the Support Vector Machine achieved the highest recall rate, reaching 98%. Additionally, a comparison of the AUC scores for the five classifiers is presented in both Table 3 and Figure 6.

Table 3: AUC scores for the five classifiers

Algorithms	AUC Score
Logistic Regression	98
K-Nearest Neighbors	99
Support Vector Machine	100
Random Forest Classifier	100
Gradient Boosting Classifier	100

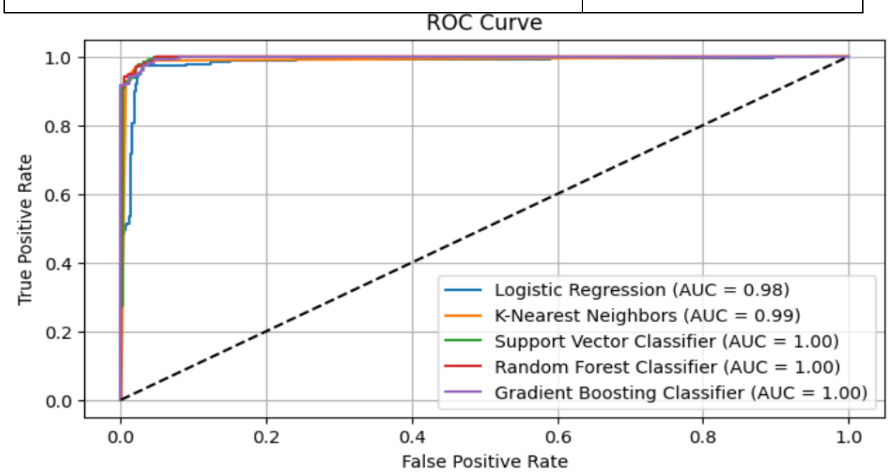

Figure 6: ROC Curve showing the performance of the models based on AUC scores

Figure 6 illustrates the Receiver Operating Characteristic (ROC) curves for all classifiers evaluated in the study, effectively demonstrating the relationship between the true positive rate and the false

December, Volume 10, Number 3, Pages 409 - 417 http://www.ijbst.fuotuoke.edu.ng /
https://doi.org/ 10.5281/zenodo.14510094 **ISSN 2488-8648**

positive rate across various thresholds. The area under the ROC curve (AUC) quantifies the overall performance of the classifiers, with higher values indicating better predictive ability. Notably, the Support Vector Machine (SVM) and the Gradient Boosting Classifier achieved exceptional AUC scores of 100%, signifying their superiority over the other classifiers tested in this study. This performance suggests that these classifiers can effectively distinguish between positive and negative instances in the dataset with higher accuracy.

Conclusion

Early detection of gestational diabetes mellitus is crucial for improving health outcomes for both mothers and their babies. Detecting the condition early can significantly slow its progression and help lower the associated mortality rate through timely and effective treatment interventions. Based on the results of this comprehensive experiment, the Support Vector Machine (SVM) model proposed in this study shows promise as a reliable clinical decision support system. This model could be utilized by healthcare providers to predict the risk of gestational diabetes mellitus in pregnant women, allowing for targeted screening and proactive management to enhance maternal and fetal health.

References

Frontino G, Succurro E, Corcoy R, Scialabba F, Poloniato A. and Scavini M. (2023). Editorial: Current and future trends in gestational diabetes diagnosis, care and neonatal outcomes. Front Endocrinol: 10.3389/fendo.2023.1270472. PMID: 38098867; PMCID: PMC10720917.

Jader, R., and Aminifar, S. (2022). Predictive model for diagnosis of gestational diabetes in the kurdistan region by a combination of clustering and classification algorithms: an ensemble approach. *Applied Computational Intelligence and Soft Computing*, *2022*(1), 9749579.

Liao, L. D., Ferrara, A., Greenberg, M. B., Ngo, A. L., Feng, J., Zhang, Z., and Zhu, Y. (2022). Development and validation of prediction models for gestational diabetes treatment modality using supervised machine learning: a population-based cohort study. *BMC medicine*, *20*(1), 307.

Mba I.N, Gav T.A, Myke-Mbata B.K, Swende T.Z, and Adebisi S.A. (2023) Rising prevalence of gestational diabetes mellitus and its associated risk factors in Makurdi, North-Central Region of Nigeria. *Afr Health Sci.* 2023 Dec;23(4):348-355. doi: 10.4314/ahs.v23i4.37.

Rani, K. J. (2020). Diabetes prediction using machine learning. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, *6*, 294-305.

Trivedi, N. K., Gautam, V., Sharma, H., Anand, A., and Agarwal, S. (2022). Diabetes prediction using different machine learning techniques. In *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)* (pp. 2173-2177). IEEE.
[[
Wu, Y. T., Zhang, C. J., Mol, B. W., Kawai, A., Li, C., Chen, L., and Huang, H. F. (2021). Early prediction of gestational diabetes mellitus in the Chinese population via advanced machine learning. *The Journal of Clinical Endocrinology & Metabolism*, *106*(3), e1191-e1205.

Xiong, Y., Zhou, Q., Wu, J., Li, X., and Xiao, X. (2020). Comparison of machine learning methods and conventional logistic regressions for predicting gestational diabetes using routine clinical data: a retrospective cohort study. *Journal of diabetes research*, *2020*(1), 4168340.

Zhang, J. and Wang, F.(2022). Prediction of gestational diabetes mellitus under cascade and ensemble learning algorithm. Comput Intell Neurosci. 2022;2022:3212738. https://doi.org/10.1155/2022/3212738. Published 2022 Jul 14

Cleveland Clinic (2024). https://my.clevelandclinic.org/health/diseases/9012-gestational-diabetes

Srivastava, Y., Khanna, P., and Kumar, S. (2019). Estimation of gestational diabetes mellitus using azure ai services. In *2019 Amity International Conference on Artificial Intelligence (AICAI)* (pp. 321-326). IEEE.

Hou, F., Cheng, Z., Kang, L., and Zheng, W. (2020, October). Prediction of gestational diabetes based on lightgbm. In *Proceedings of the 2020 Conference on Artificial Intelligence and Healthcare* (pp. 161-165).