Bigdeli *et al. BMC Medical Informatics and Decision Making* (2025) 25:3
https://doi.org/10.1186/s12911-024-02799-3

BMC Medical Informatics and Decision Making

RESEARCH Open Access

Predicting Gestational Diabetes Mellitus in the first trimester using machine learning algorithms: a cross-sectional study at a hospital fertility health center in Iran

Somayeh Kianian Bigdeli[1], Marjan Ghazisaedi[1*], Seyed Mohammad Ayyoubzadeh[1*], Sedigheh Hantoushzadeh[2] and Marjan Ahmadi[3]

Abstract

Background Gestational Diabetes Mellitus (GDM) is a common complication during pregnancy. Late diagnosis can have significant implications for both the mother and the fetus. This research aims to create an early prediction model for GDM in the first trimester of pregnancy. This model will help obstetricians and gynecologists make appropriate decisions for treating and preventing GDM complications.

Methods This applied descriptive study was conducted at the fertility health center of Vali-e-Asr Hospital in Tehran, Iran. The dataset was collected from the records of pregnant women registered in the hospital's system from 2020 to 2022. Risk factors for designing predictive models were identified through literature review, expert opinions, and clinical specialists' input. The extracted information underwent preprocessing, and six machine learning (ML) methods were developed and evaluated for GDM prediction in the first trimester of pregnancy: decision tree (DT), multilayer perceptron (MLP), k-nearest neighbors (KNN), Naïve Bayes (NB), random forest (RF), and extreme gradient boosting (XGBoost).

Results Models were evaluated using accuracy, precision, sensitivity, and the area under the receiver operating characteristic curve (AUC). Based on the glucose tolerance test (GTT) results, the RF model achieved the best performance in GDM prediction, with 89% accuracy, 86% precision, 92% recall, and 94% AUC, using demographic variables, medical history, and clinical findings. In modeling based on insulin consumption, the RF model achieved the best results with 62% accuracy, 60% precision, 63% recall, and 63% AUC in predicting GDM using demographic variables and medical history.

Conclusion The results of this study demonstrate that ML algorithms, especially RF, have acceptable accuracy in the early prediction of GDM during the first trimester of pregnancy.

*Correspondence:
Marjan Ghazisaedi
ghazimar@tums.ac.ir
Seyed Mohammad Ayyoubzadeh
smayyoubzadeh@tums.ac.ir

Full list of author information is available at the end of the article

Keywords Artificial intelligence, Gestational diabetes mellitus, Machine learning, Random forest, First trimester of pregnancy, Prediction

Introduction

Worldwide, approximately 422 million individuals have diabetes, with the majority residing in low- and middle-income countries [1–3]. Furthermore, 1.5 million deaths are directly linked to diabetes each year [4]. Gestational Diabetes Mellitus (GDM) is defined as glucose intolerance of any severity first detected during pregnancy, typically between 24 and 28 weeks gestation [5–7]. Up to 15% of expectant mothers globally experience GDM, making it a prevalent pregnancy complication [8]. It results in adverse pregnancy outcomes such as postpartum hemorrhage, infections, premature birth, large birth weight babies, and respiratory issues in newborns [9–11]. Moreover, GDM can lead to long-lasting health effects, including a higher likelihood of developing type 2 diabetes (T2DM) and cardiovascular disease (CVD) in the mother, as well as an increased risk of obesity, CVD, T2DM, and GDM in the offspring [12, 13].

The American Diabetes Association (ADA) and the International Association of Diabetes and Pregnancy Study Groups (IADPSG) recommend using a 2-hour fasting 75 g oral glucose tolerance test between 24 and 28 weeks of gestation [7, 14]. While the single-stage IADPSG approach has the benefit of requiring only one test and measuring elevated glucose levels, concerns have been raised about the overdiagnosis of GDM [6]. Earlier research has shown that diagnosing GDM at 24–28 weeks' gestation may be too late for intervention, as abnormal fetal growth can occur before the diagnosis of GDM. Thus, early screening and diagnosis of GDM in pregnancy, along with implementing suitable management, may be essential, particularly for older and obese mothers at high risk of GDM, to prevent fetal abdominal obesity [15]. Examples of abnormal growth include smaller fetuses at 24 weeks of gestation and higher growth rates in abdominal circumference compared to the non-GDM group [16].

Current clinical practices typically diagnose GDM in the second trimester, between 24 and 28 weeks of gestation. However, this timing may be too late for effective intervention, as fetal growth abnormalities can occur before diagnosis. Our study aims to address this gap by developing a model for earlier prediction of GDM in the first trimester.

In today's world, the increasing amount of information and the intricacy of decision-making have made it more crucial to utilize artificial intelligence (AI) systems to assist in medical decision-making [17]. There are roughly two categories of AI techniques: symbolic AI and computational intelligence. Symbolic AI focuses on creating knowledge-based systems, whereas computational intelligence concentrates on developing computational methods inspired by nature [18]. Data Mining (DM) and Machine Learning (ML) platforms belong to the field of AI and utilize methods and approaches for extracting information and insights from extensive datasets. They use algorithms to create and extract new knowledge through association, classification, prediction, and clustering [19].

While our study uses established data mining techniques, it contributes to the field in several ways. First, we focus on first-trimester prediction of GDM, which is relatively new and clinically relevant. Second, this is one of the first studies to apply machine learning techniques for GDM prediction in an Iranian population. Finally, our comprehensive comparison of various machine learning algorithms provides insight into their relative performances for this specific prediction task.

Various studies have explored using DM and ML in managing and analyzing patient records and disease biomarkers. Data mining classifiers [20] are commonly utilized to predict GDM, and various ML algorithms have been employed for this purpose [16, 21–25]. This research was conducted to meet the clinical requirements of the fertility health center at Vali-e-Asr Hospital in Iran. The objective was to create an ML model for early prediction of GDM in the first trimester of pregnancy.

Materials and methods

This research employs a practical, descriptive-developmental approach and was conducted in four primary phases. Figure 1 illustrates the general process of this study. We utilized the Python programming language to develop machine learning models, NumPy for data preprocessing, and the Scikit-learn library to create the models.

Dataset description and participants

The dataset for this study was derived from the Reproductive Health Center of Vali-e-Asr Hospital. The study data were obtained from the records of pregnant women registered in the hospital's system between 2020 and 2022. A total of 16,730 records were extracted from the system and subsequently entered into a CSV file. The Medical Ethics Committee of the first author's university approved the study (Ethical approval number: IR.TUMS.SPH.REC.1402.039).

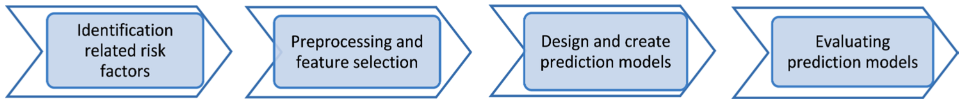

Fig. 1 General stages of the current research

Entry and exit criteria for patients

Inclusion criteria

1. First trimester of pregnancy
2. Complete baseline data

Exclusion criteria

3. Pre-existing diabetes
4. History of GDM
5. Incomplete follow-up data

Our initial dataset consisted of 16,730 records. However, after applying our inclusion criteria and removing records with missing data, our final sample size was reduced to 743 records for insulin consumption modeling and 106 records for GTT modeling. This significant reduction was primarily due to the strict requirement of having complete data for all relevant variables in the first trimester. While this reduction limits the generalizability of our findings, we believe the resulting dataset still provides valuable insights into early GDM prediction.

Identifying predictive variable related to GDM and patient selection

The research began with a comprehensive literature review of early GDM prediction articles from scientific databases including PubMed, Scopus, Web of Science, and Embase to identify relevant risk factors. These risk factors were then validated by a clinical expert.

A checklist was developed, comprising questions about the participating experts, 39 identified risk factors, and suggested review items. This checklist used a 5-point Likert scale and was shared electronically with experts at the Vali-e-Asr Reproductive Health Clinical Research Center. Eighteen experts participated in completing this checklist.

After compilation, the findings were analyzed using Excel. Based on clinical expert opinion, a maximum acceptable threshold of 3 was established, with risk factors scoring 3 or higher identified as significant indicators. The experts also recommended additional items for consideration, such as complete history records, first-trimester vitamin D3 levels, and current obstetric status, which were subsequently reviewed and approved in a specialized meeting.

Furthermore, the 'Result of NT sonographic' was examined by clinical experts to determine its impact on the outcome. Based on its assessed effect, this factor was incorporated into the modeling process.

We included 'NT sonographic' results in our model based on expert evaluation of its potential impact on GDM prediction. While not traditionally associated with GDM, recent research suggests that first-trimester ultrasound markers may have predictive value for various pregnancy complications, including GDM.

Preprocessing

The research center's database provided the necessary information based on approved risk factors. This data was subsequently processed through data mining, with careful consideration given to the handling of missing data. Specifically:

6. For risk factors with less than 30% missing numeric values, adjustments were made using multiple imputation [26].
7. Non-numeric values with missing data were categorized based on their frequency.

Outlier data was identified using the first to third quartile interval and then adjusted to align with the median's lower and upper limits [27]. In the final stage, numerical risk factors were normalized using the Min-Max Scaler method [10] to ensure accurate comparison based on their existence and categories, and placed within the 0 to 1 range.

Following this, the datasets were balanced using the Synthetic Minority Over-sampling Technique (SMOTE) method [12] due to the imbalance in the number of samples for each class (positive GDM diagnosis and negative diagnosis of GDM).

Based on the clinical consultant's recommendation, all risk factors were included in the modeling process to identify the effective ones.

Feature selection

In data mining operations, a key challenge is identifying the connections between features within a dataset and the outcomes. Selecting the right features is crucial for successful data mining, particularly in scenarios involving numerous features and dataset variations. Feature selection involves identifying important variables and excluding those that are irrelevant or redundant, as defined by [28] Various fields employ feature selection to remove irrelevant or duplicate features from their applications [29].

In this research, the initial step involved reviewing and validating the set of identified risk factors, first by a clinical specialist. Following the review of the files, the contents were compared with the confirmed risk factors. Subsequently, a questionnaire was compiled, comprising:

1. Inquiries about the general details of the participating experts
2. 39 recognized risk factors
3. Suggested review items

Creating prediction models

At this stage, to predict GDM early, two modeling experiments were considered based on clinical expert opinion:

1. Modeling according to the diagnosis of GDM based on insulin consumption
2. Modeling according to the diagnosis of GDM based on the result of the GTT test

Six machine learning models were designed and created using the processed data after adjusting the hyperparameters:

1. Decision Tree (DT)
2. Multilayer Perceptron (MLP)
3. K-Nearest Neighbors (KNN)
4. Naïve Bayes (NB)
5. Random Forest (RF)
6. Extreme Gradient Boosting (XGBoost)

We applied the grid search method for hyperparameter tuning [30]. Scikit-Learn's GridSearchCV class is used for this purpose. It assesses every possible combination of parameter values and selects the optimal set of parameters.

Model evaluation

During this stage, the constructed models underwent evaluation based on multiple criteria:

1. Accuracy
2. Other predictive algorithm evaluation metrics [31]

To conduct a more comprehensive comparison of algorithm performances, we plotted the Area Under the Receiver Operating Characteristic curve (AUC) for each model in a standardized format. This approach facilitated the identification of the best prediction model [32].

To ensure model robustness, we implemented 5-fold cross-validation for all models. This approach helps verify the models' performance across different subsets of our data, providing a more reliable estimate of their generalizability [33].

Results

Characteristics of patients in GDM prediction based on insulin consumption

For the initial modeling experiment, we defined our target variables for prediction as follows:

- Positive class: Insulin use ($n = 335$)
- Negative class: Non-insulin use with a confirmed negative diabetes outcome in the second trimester ($n = 408$)

After preprocessing the data to address missing and outlier values, we loaded 743 records into the Jupyter Notebook environment. This dataset comprised:

- 335 records with GDM
- 408 records without GDM
- 11 variables for preprocessing

We utilized the Wrapper Forward feature selection technique to identify key variables. This approach highlighted three important variables:

1. Age
2. BMI (Body Mass Index)
3. History of Abortion in Previous Obstetrics (number of abortions in previous pregnancies)

After evaluating the significant variables in the modeling results and consulting with a clinical expert, we decided to include all variables as input for the model (Table 1).

Characteristics of patients in GDM prediction using the GTT diagnostic test

In the second experiment, our goal was to model and predict GDM based on the results of the GTT (Glucose Tolerance Test) diagnostic test in the second trimester. The target variables for prediction were defined as follows:

- Positive class: Positive GTT test results ($n = 149$)
- Negative class: Negative GTT test results ($n = 537$)

It's worth noting that 16,043 records were excluded from the analysis.

Based on the clinical expert's evaluation, the dataset was refined to incorporate only pertinent variables, resulting in 67 records:

- 14 records associated with positive GDM (class 1)
- 53 records linked to negative GDM (class 0)

Table 1 Variables included in modeling based on insulin consumption

Variables included in modeling		
Variable name	**Type**	**Values**
Age	Quantitative	19–48 (year)
BMI	Quantitative	16–169 (kg/m2)
History of hypertension	Categorical	Yes No
Family history of diabetes mellitus	Categorical	Yes No
Family history of hypertension	Categorical	Yes No
History of abortion in previous pregnancies	Quantitative	0–6
History of cardiovascular diseases	Categorical	Yes No
History of endocrine metabolic disease	Categorical	Yes No
History of digestive disease	Categorical	Yes No
Current obstetric normal	Categorical	Yes No
History of infertility	Categorical	Yes No

Table 2 Variables included in modeling based on GTT test

Variables included in modeling		
Variable	**Type**	**Values**
Age	Quantitative	20–48 (year)
BMI	Quantitative	17–51 (kg/m2)
History of hypertension	Categorical	Yes No
History of cardiovascular diseases	Categorical	Yes No
History of abortion in previous pregnancies	Quantitative	0–4
First trimester FBS	Quantitative	62–110
First trimester HB	Quantitative	10.1–15.7
First trimester Hct	Quantitative	2-45.4
First trimester Cr	Quantitative	0-1.8
First trimester PLT	Quantitative	150-336000
First trimester vit D3	Quantitative	6–60
First trimester NT Sonographic (nt)	Quantitative	1-3.2
First trimester NT Sonographic (crl)	Quantitative	45.2–78

All preprocessing steps for this dataset were completed.

In this modeling approach, the key variables identified include:

1. Age
2. BMI
3. History of Abortion in Previous Obstetrics
4. Fasting Blood Sugar (FBS) in the first trimester

After evaluating the results of including these variables in the modeling and consulting with the clinical expert, we decided to include all variables as input for the model (Table 2).

Modeling according to the diagnosis of GDM based on insulin consumption

The evaluation results for the forecasting models are illustrated in Fig. 2. According to this graph, the RF model outperformed other models across various metrics:

- Mean Accuracy: 62%
- Mean Precision: 60%
- Mean Recall: 63%
- Mean F1-score: 63%

Furthermore, as shown in Fig. 3, the RF model demonstrates a higher AUC compared to other models.

The confusion matrix results for the RF model are presented in Fig. 4. The values for each category in the confusion matrix are as follows:

- True Positives (TP): 255
- True Negatives (TN): 237
- False Negatives (FN): 153
- False Positives (FP): 171

Modeling considering the diagnosis of GDM using the GTT diagnostic test

The evaluation results for the forecasting models are illustrated in Fig. 5. According to this graph, the RF model outperformed other models across various metrics:

- Mean Accuracy: 89%
- Mean Precision: 86%
- Mean Recall: 92%
- Mean F1-score: 89%

Furthermore, as shown in Fig. 6, the RF model demonstrates a higher AUC compared to other models.

The results related to the confusion matrix of the RF model are shown in Fig. 7. The value of TP for this model was 49, TN equal to 43, FN equal to 4 and FP equal to 10.

Comparison of insulin consumption and GTT models

While both the insulin consumption and GTT models demonstrated good predictive performance, there were notable differences. The GTT-based model achieved higher accuracy (89% vs. 62%) and AUC (94% vs. 64%) compared to the insulin consumption model. This discrepancy may be due to the GTT being a more direct measure of glucose intolerance. However, the insulin consumption model has the advantage of being

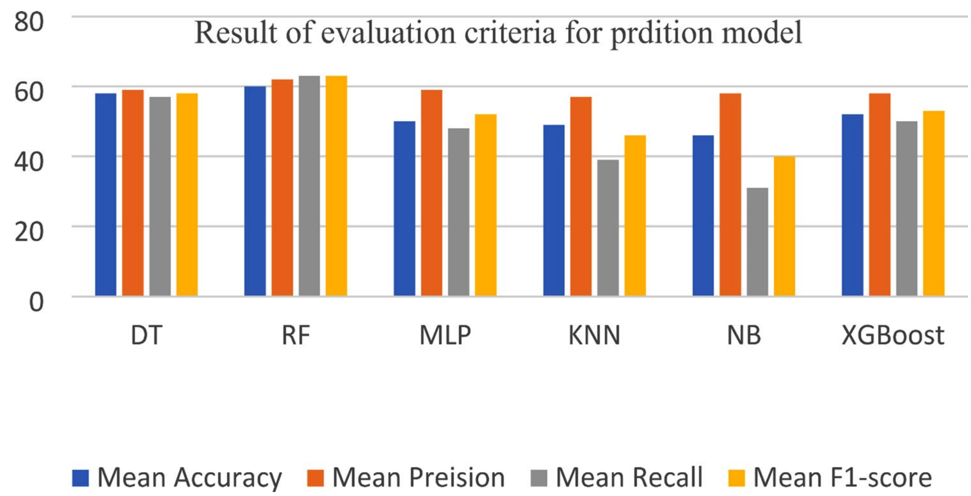

Fig. 2 The results of evaluation criteria of prediction models in modeling based on insulin consumption

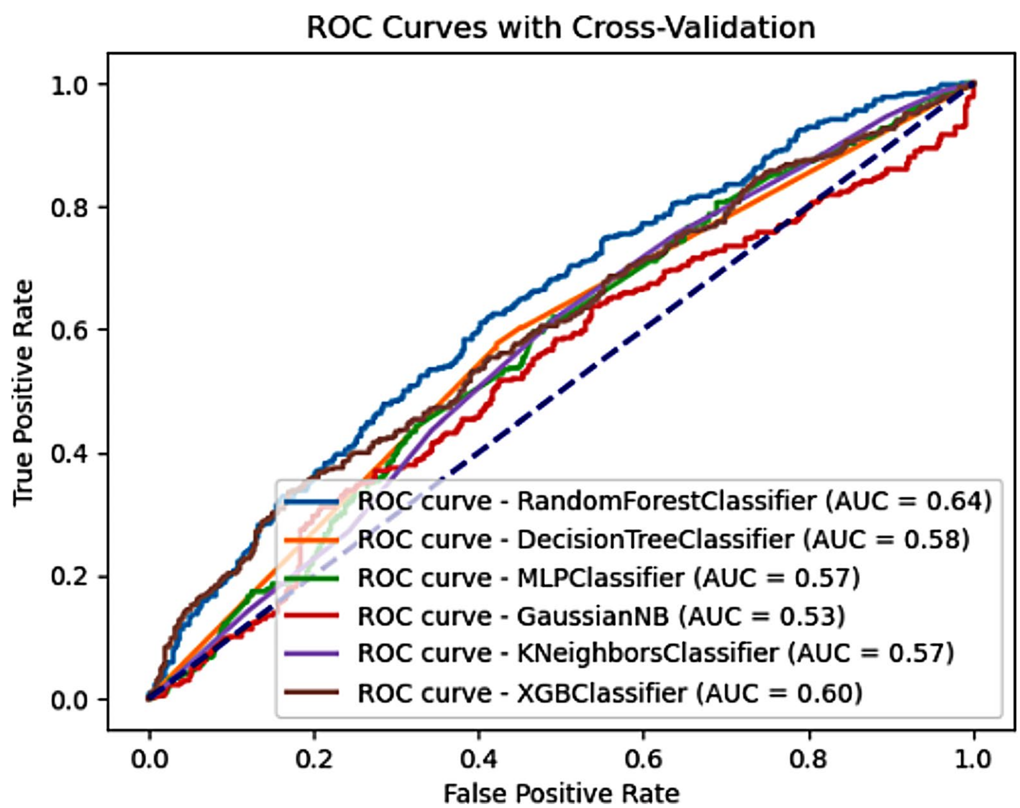

Fig. 3 The AUC of predictive models in modeling based on insulin consumption

applicable to a larger population, as it doesn't require the GTT test results.

Discussion

The present study focused on developing an early prediction model for GDM during the first trimester of pregnancy using machine learning algorithms. To achieve this, we developed DT, RF, k-NN, MLP, XGBoost, and NB models using a dataset comprising demographic variables, medical history, and clinical findings from pregnant women.

Our experimental results showed that in modeling based on the GTT, RF had the best performance among the other ML techniques, with 89% accuracy, 86% precision, 92% recall, and 94% AUC. In modeling based on insulin consumption, the RF model again demonstrated the best performance, achieving 62% accuracy, 60% precision, 63% recall, and an AUC of 64%.

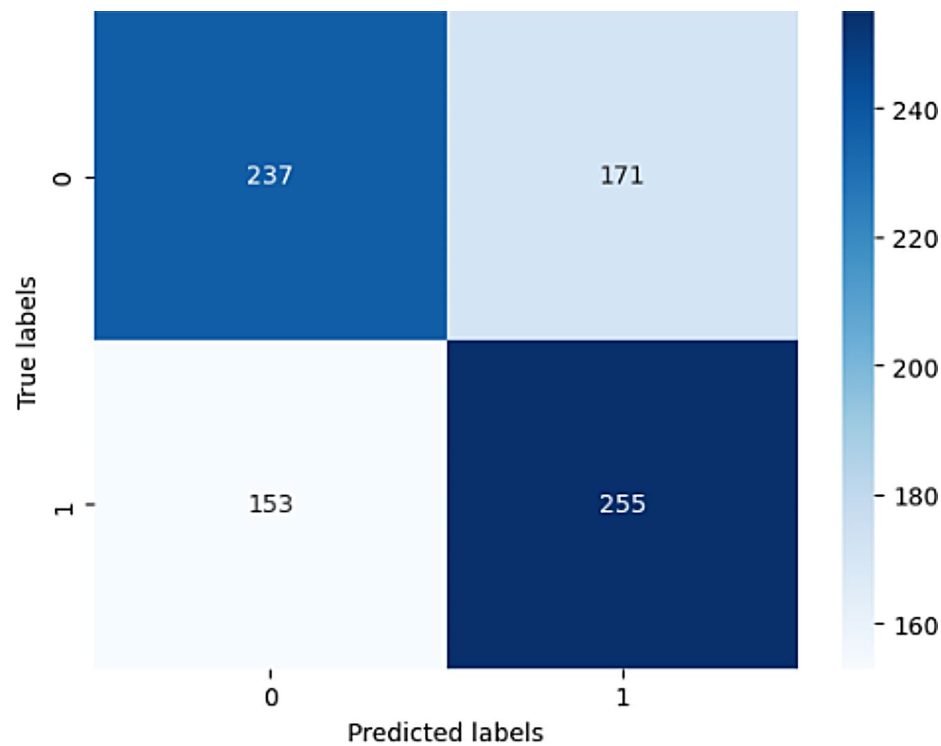

Fig. 4 Values related to the RF model confusion matrix based on insulin consumption

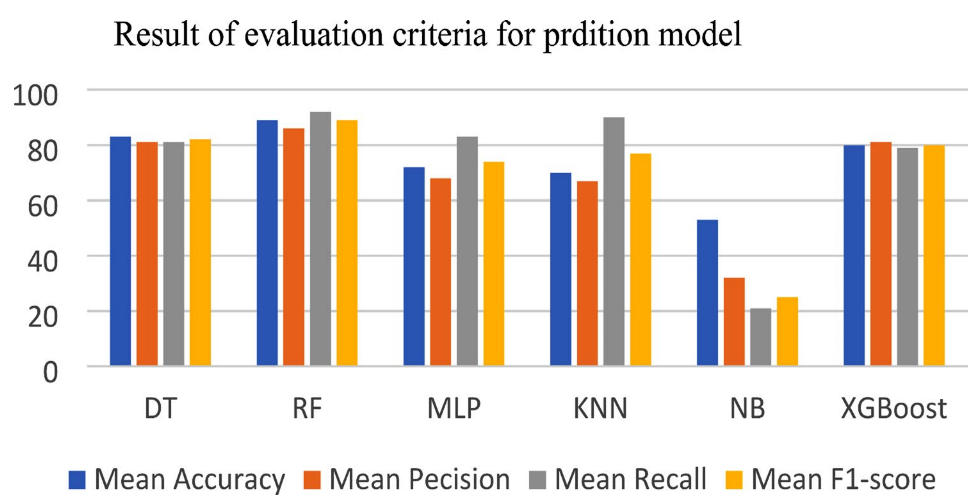

Fig. 5 The results of evaluation criteria of prediction models in modeling based on the result of the GTT test

The modeling based on insulin consumption utilized 11 variables in this study, while the modeling based on the GTT test utilized 13 variables, as advised by experts. Janak et al. reviewed 25 studies to identify the most significant variables. Their results indicated that maternal age, BMI, FBS, and family history of diabetes were the most commonly utilized variables for predicting GDM [31]. Zarei et al. used five variables including FBS at the first pregnancy examination, history of GDM in previous pregnancies, BMI, mother's age, and family history of diabetes, which had the greatest impact on their modeling [34]. Rezaei et al., in Kermanshah City, used demographic data, pregnancy rate, diabetes, fertility parameters, and test results [35]. However, several studies have attempted to use unusual risk factors like genetic factors in modeling [23, 36, 37].

Considering the process of diagnosis and screening of pregnant women in the first trimester, modeling using demographic information, medical records, routine tests, and first-trimester ultrasound results appears to be more effective in real-world environments. Based on our findings using wrapper forward feature selection, age, BMI,

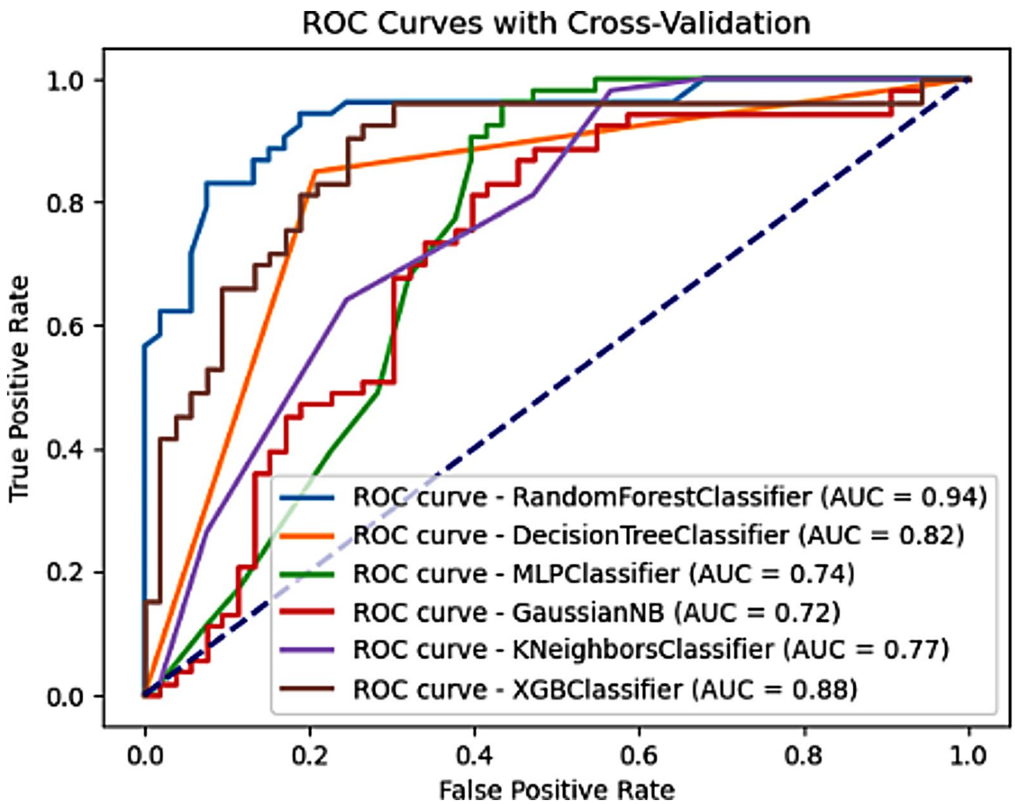

Fig. 6 The AUC of predictive models in modeling based on the result of the GTT test

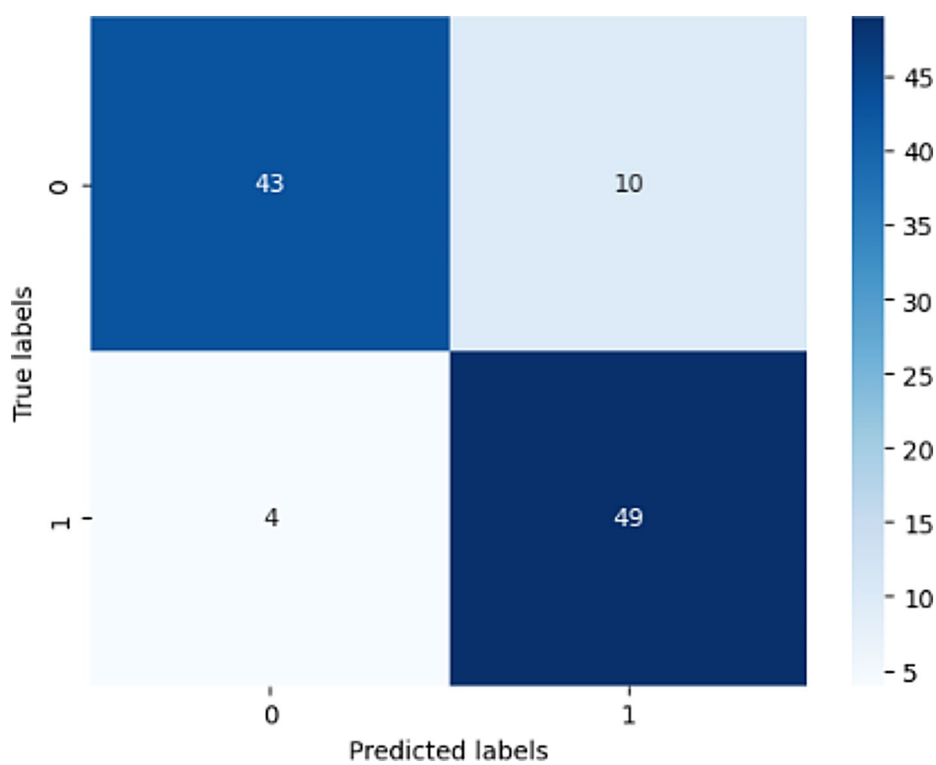

Fig. 7 Values related to the RF model confusion matrix based on the result of the GTT test

and history of abortion in previous pregnancies were important variables in the insulin consumption-based modeling. In the second modeling approach, these three variables plus first-trimester FBS had a greater impact on the model than other variables.

Various studies have evaluated the application of ML techniques in predicting GDM [23, 36, 38, 39]. Sumathi et al. assessed the performance of ML algorithms for predicting early GDM. In their dataset, the RF model achieved the highest performance in predicting GDM, with an accuracy of 77% [22]. Jader et al. used classification methods including DT, NB, RF, and KNN, achieving 92% accuracy, while Support Vector Machine (SVM) and Logistic Regression (LR) achieved 90% accuracy for group techniques [24]Yan Ting et al. developed an LR model using seven variables that reached 77% AUC, identifying it as the optimal model for clinical centers [16]. Rezaei et al. reported a perceptron Neural Network (NN) model with 79% AUC, 83% accuracy, 62% sensitivity, and 95% specificity as their proposed model [35]. Zarei et al.'s research led to the development of a predictive model for GDM using a combination of the DT model and Artificial Neural Network (ANN). This model achieved an AUC of 86% and a sensitivity of 92.1% [34].

The implications of our findings are significant for clinical practice. Early prediction of GDM can lead to earlier interventions, potentially improving maternal and fetal outcomes. In resource-constrained settings, early identification of high-risk patients can help prioritize care and interventions. Moreover, our models contribute to the growing field of personalized medicine in obstetrics, offering a tool for individualized risk assessment and management strategies for pregnant women.

Limitations

This study had several limitations that warrant consideration:

1) Small sample size in modeling which can limit the statistical power of our various analyses. This consequence may lead to: (a) The reduced capability to detect significant effects may fail to detect important associations. (b) Greater chance of Type II errors (false negatives) (c) Less precise estimates of effect size, as reflected by wider confidence intervals. (d) Risk of overfitting of the prediction models at the cost of generalizability
2) Single-center design: Data collected within a single healthcare facility might not be wholly representative of the general population. This might, therefore, lead to the following setbacks: (a) Bias towards the particular patient demographics this center serves. (b) Effects of institution-specific practices or protocols on the results (c) Limited variability in environmental or socio-economic factors that may influence pregnancy outcomes
3) Lack of External Validation: Without validation in an independent dataset, the model's performance in other settings remains uncertain. Consequently: (a) The predictive accuracy might not be uniform across populations or health systems. (b) The model may be overfitting to the characteristics in our study population. (c) Clinical utility of the model in real-world scenarios outside of our center remains to be seen

Despite these limitations, we used variables that can be easily and routinely measured in the first trimester without imposing significant costs on the patients or the healthcare system. This, therefore, increases the likelihood of practical implementation if the model is validated by future studies.

Future directions

The future research would focus on a number of aspects: first, multi-center validation studies-inherent in the meaning of enlargement of applicability of our results. Collaborative studies across various healthcare centers of Iran should be formulated and actually implemented by standard protocols for collecting data. This would mean creating a common infrastructure of the database, early anonymization of patients to enable performance comparisons of models across demographics and geographies while considering different clinical practices in various demographic regions.

The model development has to be improved, focusing on the use of other ML algorithms, including Deep Learning and ensemble methods, and developing hybrid models which can integrate several algorithms. This would mean using automatic feature selection methods, applying updating of models in real time, and verifying whether methods of transfer learning might improve the accuracy.

Other key components of future work would include clinical implementation research, including user-friendly interface development for use by clinical staff and usability studies with healthcare providers. This also encompasses the analysis of the impact on clinical decision-making, the assessment of the cost-effectiveness of implementation, and the monitoring of patient outcomes after model deployment.

Different variables will be integrated that involve expansion, which could enhance the performance of the model. This includes the integration of genetic markers and biomarkers, life course and environmental exposure variables, socioeconomic indicators, pregnancy-specific hormonal markers, and the influence/importance of medication history on the accuracy of the prediction.

Ultimately, longitudinal studies of the implications of our predictive model for the long term will be obtained by the design of cohort studies with extended follow-up and long-term maternal and fetal outcomes. Successive pregnancy tests of the predictive value of the model should be considered in the studies that make early interventions based on model predictions and monitor the incidence of type 2 diabetes among GDM patients. The results of such extensive studies will add inestimable insight into the predictive model's long-term efficiency and clinical utility.

Conclusion

The results of this study suggest that ML algorithms, especially RF, can be effectively utilized to predict GDM during the first trimester of pregnancy with high accuracy, particularly when trained on demographic, medical history, and clinical data. Our model demonstrated acceptable performance in GTT and insulin consumption analysis, indicating that RF could serve as a robust tool for early GDM prediction. These results align with previous studies in ML applications for obstetric risk assessment, suggesting that RF has a consistent edge over other algorithms in similar clinical prediction tasks. The use of ML-based prediction models can improve the management of GDM, although these models are currently more commonly used in clinical diagnosis.

Based on the experiences of several medical centers worldwide in utilizing these algorithms in medical decision support systems, we recommend incorporating this model into electronic health records. This integration could potentially enhance early detection and management of GDM, leading to improved patient outcomes.

Future works should aim to address the model's current limitations and establish a broader foundation for the use of AI-driven prediction tools in prenatal healthcare. With rigorous validation, these models could offer an invaluable addition to medical decision support, ultimately contributing to the goal of personalized and preventative care in pregnancy.

Abbreviations

GDM	Gestational Diabetes Mellitus
T2DM	Type 2 Diabetes Mellitus
CVD	Cardiovascular Disease
ADA	American Diabetes Association
IADPSG	International Association of Diabetes and Pregnancy Study Groups
AI	Artificial Intelligence
DM	Data Mining
ML	Machine Learning
DT	Decision Tree
MLP	Multilayer Perceptron
KNN	K-Nearest Neighbors
NB	Naïve Bayes
RF	Random Forest
AUC	Area Under the Receiver Operating Characteristic Curve
BMI	Body Mass Index
FBS	Fasting Blood Sugar
GTT	Glucose Tolerance Test
TP	True Positive
TN	True Negative
FN	False Negative
FP	False Positive
LR	Logistic Regression
SVM	Support Vector Machine
NN	Neural Network
ANN	Artificial Neural Network

Acknowledgements
Not applicable.

Author contributions
M.A. and S.H. conceptualized the work. S.M.A. contributed to the methodology and reviewing the work. M.Gh. supervised the work and reviewed the work. S.K.B. wrote the original draft, conducted the data curation, and performed the formal analysis.

Funding
Not applicable.

Data availability
The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
This retrospective study analyzed datasets covering the period from 2020 to 2023. The Research Ethics Committee of the School of Public Health & Allied Medical Sciences, Tehran University of Medical Sciences, approved the study (approval ID: IR.TUMS.SPH.REC.1402.039). Due to the retrospective nature of the study and the use of existing anonymized data, the ethics committee waived the requirement for individual informed consent, in accordance with institutional and national guidelines. The ethical approval certificate can be accessed at: https://ethics.research.ac.ir/ProposalCertificateEn.php?id=331317.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Health Information Management, School of Allied Medical Sciences, Tehran University of Medical Sciences, Tehran, Iran
[2]Vali-E-Asr Reproductive Health Research Center, Family Health Research Institute, Imam Khomeini Hospital Complex, Tehran University of Medical Sciences, Tehran, Iran
[3]Department of Obstetrics and Gynecology, Tehran University of Medical Sciences, Tehran, Iran

Received: 22 August 2024 / Accepted: 2 December 2024
Published online: 03 January 2025

References

1. Zhu Y, Zhang C. Prevalence of gestational diabetes and risk of progression to type 2 diabetes: a global perspective. Curr Diab Rep. 2016;16(1):7.
2. Shakeri E, Crump T, Weis E, Souza R, Far B. Using SHAP analysis to detect Areas contributing to Diabetic Retinopathy Detection. In 2022 IEEE 23rd International Conference on Information Reuse and Integration for Data Science (IRI). 2022. 166–71 p.
3. International Diabetes Federation [Internet]. 2021. IDF Diabetes Atlas. Available from: https://diabetesatlas.org/.
4. World Health Organization. Diabetes [Internet]. 2023. https://www.who.int/health-topics/diabetes#tab=tab_1
5. Hod M, Kapur A, Sacks DA, Hadar E, Agarwal M, Di Renzo GC, et al. The International Federation of Gynecology and Obstetrics (FIGO) Initiative on

gestational diabetes mellitus: a pragmatic guide for diagnosis, management, and care. Int J Gynaecol Obstet off Organ Int Fed Gynaecol Obstet. 2015;131:S173–211.

6. Fu J, Retnakaran R. The life course perspective of gestational diabetes: an opportunity for the prevention of diabetes and heart disease in women. EClinicalMedicine. 2022;45:101294.
7. Association AD. Gestational Diabetes Mellitus. Diabetes Care [Internet]. 2003;26(suppl_1):s103–5. https://doi.org/10.2337/diacare.26.2007.S103
8. Ogurtsova K, da Rocha Fernandes JD, Huang Y, Linnenkamp U, Guariguata L, Cho NH, et al. IDF Diabetes Atlas: global estimates for the prevalence of diabetes for 2015 and 2040. Diabetes Res Clin Pract. 2017;128:40–50.
9. Agarwal P, Morriseau TS, Kereliuk SM, Doucette CA, Wicklow BA, Dolinsky VW. Maternal obesity, diabetes during pregnancy and epigenetic mechanisms that influence the developmental origins of cardiometabolic disease in the offspring. Crit Rev Clin Lab Sci. 2018;55(2):71–101.
10. Tinker SC, Gilboa SM, Moore CA, Waller DK, Simeone RM, Kim SY, et al. Specific birth defects in pregnancies of women with diabetes: National Birth defects Prevention Study, 1997–2011. Am J Obstet Gynecol. 2020;222(2):176. e1-176. e11.
11. Zhang C, Olsen SF, Hinkle SN, Gore-Langton RE, Vaag A, Grunnet LG, et al. Diabetes & women's Health (DWH) study: an observational study of long-term health consequences of gestational diabetes, their determinants and underlying mechanisms in the USA and Denmark. BMJ Open. 2019;9(4):e025517.
12. Zhang X, Zhao X, Huo L, Yuan N, Sun J, Du J et al. Risk prediction model of gestational diabetes mellitus based on nomogram in a Chinese population cohort study. Sci Rep [Internet]. 2020;10(1):21223. https://doi.org/10.1038/s41598-020-78164-x
13. Plows JF, Stanley JL, Baker PN, Reynolds CM, Vickers MH. The pathophysiology of gestational diabetes Mellitus. Int J Mol Sci. 2018;19(11).
14. Metzger BE, Gabbe SG, Persson B, Buchanan TA, Catalano PA, Damm P, et al. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. Diabetes Care. 2010;33(3):676–82.
15. Kim W, Park SK, Kim YL. Gestational diabetes mellitus diagnosed at 24 to 28 weeks of gestation in older and obese women: is it too late? PLoS ONE. 2019;14(12):e0225955.
16. Wu YT, Zhang CJ, Mol BW, Kawai A, Li C, Chen L, et al. Early Prediction of Gestational Diabetes Mellitus in the Chinese Population via Advanced Machine Learning. J Clin Endocrinol Metab. 2021;106(3):E1191–205.
17. Sadoughi F, Sheikhtaheri A. Applications of Artificial Intelligence in Clinical Decision Making: Opportunities and Challenges. Heal Inf Manag [Internet]. 2011;8(3):445. https://www.sid.ir/FileServer/JF/76513901916.pdf
18. Karri V. Introduction of Artificial Intelligence techniques and approaches. 2021.
19. Duque J, Godinho A, Moreira J, Vasconcelos J. Data Science with Data Mining and Machine Learning A design science research approach. Procedia Comput Sci [Internet]. 2024;237:245–52. https://www.sciencedirect.com/science/article/pii/S1877050924011189
20. Pernabas JB, Fidele SF, Vaithinathan KK. Enhancing Greedy Web Proxy caching using Weighted Random Indexing based Data Mining Classifier. Egypt Informatics J [Internet]. 2019;20(2):117–30. https://www.sciencedirect.com/science/article/pii/S1110866517300841
21. Cubillos G, Monckeberg M, Plaza A, Morgan M, Estevez PA, Choolani M, et al. Development of machine learning models to predict gestational diabetes risk in the first half of pregnancy. BMC Pregnancy Childbirth. 2023;23(1):469.
22. Amarnath S, Selvamani M, Varadarajan V. Prognosis model for gestational diabetes using machine learning techniques. Sensors Mater [Internet]. 2021;33(9):3011–25. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85117411955&doi=10.18494%2FSAM.2021.3119&partnerID=40&md5=27b35c53c520f52c703e6098906bf827
23. Zhang J, Wang F. Prediction of gestational diabetes Mellitus under Cascade and Ensemble Learning Algorithm. Comput Intell Neurosci. 2022;2022:3212738.
24. Jader R, Aminifar S. Predictive Model for Diagnosis of Gestational Diabetes in the Kurdistan Region by a Combination of Clustering and Classification Algorithms: An Ensemble Approach. Appl Comput Intell Soft Comput. 2022;2022.
25. Reddy SS, Gadiraju M, Preethi NM, Rao VVRM. A Novel Approach for Prediction of Gestational Diabetes based on clinical signs and risk factors. ICST Trans Scalable Inf Syst. 2023;10(3):e8.
26. Patrician PA. Multiple imputation for missing data. Res Nurs Health. 2002;25(1):76–84.
27. Whaley DL III. The interquartile range: theory and estimation. East Tennessee State University; 2005.
28. Afrash MR, Shafiee M, Kazemi-Arpanahi H. Establishing machine learning models to predict the early risk of gastric cancer based on lifestyle factors. BMC Gastroenterol. 2023;23(1):6.
29. Dash M, Liu H. Feature selection for classification. Intell data Anal. 1997;1(1–4):131–56.
30. Alibrahim H, Ludwig SA. Hyperparameter optimization: comparing genetic algorithm against grid search and bayesian optimization. 2021 IEEE Congress on Evolutionary Computation (CEC). In: Kraków, Poland. 2021. pp. 1551–9.
31. Zhang Z, Yang L, Han W, Wu Y, Zhang L, Gao C et al. Machine Learning Prediction Models for Gestational Diabetes Mellitus: Meta-analysis. J Med Internet Res [Internet]. 2022;24(3):e26634. https://doi.org/10.2196/26634
32. Wu H, Jun Meng F, Review on Evaluation Criteria of Machine Learning Based on Big Data. J Phys Conf Ser [Internet]. 2020;1486(5):52026. https://doi.org/10.1088/1742-6596/1486/5/052026
33. Kohavi R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. 2001;14.
34. Zarei J, Izadi M, Azizi ̄, Nouhjah S. Early Prediction of Gestational Diabetes Using Decision Tree and Artificial Neural Network Algorithms. IJEM [Internet]. 2022;24(1):1–11. http://ijem.sbmu.ac.ir/article-1-3015-en.html
35. Rezaei M, Fakhri N, Rajati F, Shahsavari S. Comparison of gestational diabetes prediction with artificial neural network and decision tree models. Tehran Univ Med Sci J. 2019;77(6):359–67.
36. Jinlan G, Guanghui F, Jiequan O, Tingting W, COMBATING CLASS-IMBALANCE AND, OUTLIERS IN GESTATIONAL DIABETES MELLITUS PREDICTION. Acta Medica Mediterr [Internet]. 2022;38(2):1167–74. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85128124455&doi=10.19193%2F0393-6384_2022_2_178&partnerID=40&md5=f1e55c381cebd80c254bd23e70080384
37. Wang X, Wang Y, Zhang S, Yao L, Xu S. Analysis and Prediction of Gestational Diabetes Mellitus by the Ensemble Learning Method. Int J Comput Intell Syst [Internet]. 2022;15(1). https://www.scopus.com/inward/record.uri?eid=2-s2.0-85137149228&doi=10.1007%2Fs44196-022-00110-8&partnerID=40&md5=4a141e4aa53154afb05696227b380ea9
38. Wang J, Lv B, Chen X, Pan Y, Chen K, Zhang Y, et al. An early model to predict the risk of gestational diabetes mellitus in the absence of blood examination indexes: application in primary health care centres. BMC Pregnancy Childbirth. 2021;21(1):814.
39. Assegie TA, Suresh T, Purushothaman R, Ganesan S, Kumar NK. Early prediction of gestational diabetes with parameter-tuned K-Nearest neighbor classifier. J Robot Control. 2023;4(4):452–7.

Publisher's note