

RESEARCH

Open Access



# The early prediction of gestational diabetes mellitus by machine learning models

Yeliz Kaya<sup>1\*</sup>, Zafer Bütün<sup>2</sup>, Özer Çelik<sup>3</sup>, Ece Akça Salık<sup>4</sup>, Tuğba Tahta<sup>5</sup> and Arzu Altun Yavuz<sup>6</sup>

## Abstract

**Background** We aimed to determine the best-performing machine learning (ML)-based algorithm for predicting gestational diabetes mellitus (GDM) with sociodemographic and obstetrics features in the pre-conceptional period.

**Methods** We collected the data of pregnant women who were admitted to the obstetric clinic in the first trimester. The maternal age, body mass index, gravida, parity, previous birth weight, smoking status, the first-visit venous plasma glucose level, the family history of diabetes mellitus, and the results of an oral glucose tolerance test of the patients were evaluated. The women were categorized into groups based on having and not having a GDM diagnosis and also as being nulliparous or primiparous. 7 common ML algorithms were employed to construct the predictive model.

**Results** 97 mothers were included in the study. 19 and 26 nulliparous were with and without GDM, respectively. 29 and 23 primiparous were with and without GDM, respectively. It was found that the greatest feature importance variables were the venous plasma glucose level, maternal BMI, and the family history of diabetes mellitus. The eXtreme Gradient Boosting (XGB) Classifier had the best predictive value for the two models with the accuracy of 66.7% and 72.7%, respectively.

**Discussion** The XGB classifier model constructed with maternal sociodemographic findings and the obstetric history could be used as an early prediction model for GDM especially in low-income countries.

**Keywords** Gestational diabetes mellitus, Machine learning, Prognostic prediction model, Maternal health

## Introduction

Gestational diabetes mellitus (GDM) is a major pregnancy-related disorder that affects fetal and maternal health, causes preterm labor, and could be a risk factor in the mother developing diabetes mellitus type 2 after the delivery in the future. GDM not only affects fetomaternal health, but it can also cause psychological problems for women during and after pregnancy. This disease, which can cause major health problems, has been found in nearly 20% of pregnant women around the world and could be identified as a common public health problem [1]. Although GDM screening in the first trimester is not recommended for women [2, 3], early identification of women at high risk of developing GDM may be more important to support them at an earlier stage. Various

\*Correspondence:

Yeliz Kaya  
yelizyilmazturk@gmail.com

<sup>1</sup>Faculty of Health Sciences, Department of Gynecology and Obstetrics Nursing, Eskişehir Osmangazi University, Eskişehir, Turkey

<sup>2</sup>Hoşnudiye Mah. Aysen Sokak Dorya Rezidans, A Blok no:28/77, Eskişehir, Turkey

<sup>3</sup>Faculty of Science, Department of Mathematics-Computer Science, Eskişehir Osmangazi University, Eskişehir, Turkey

<sup>4</sup>Department of Gynecology and Obstetrics, Eskişehir City Hospital, Eskişehir, Turkey

<sup>5</sup>Ankara Medipol University, Health Services Vocational School, Ankara, Turkey

<sup>6</sup>Faculty of Science, Department of Statistics, Eskişehir Osmangazi University, Eskişehir, Turkey



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

predictive models have been developed and validated for the early detection of GDM [4–6]. Classical statistical software was used to develop these models, however, the major limitation of these is that they do not have enough predictive power to be used as a routine antenatal screening program [4]. In order to overcome these limitations, machine learning-based algorithms have been employed to develop high-performance GDM predictive models [7–9]. However, there are limited studies on artificial intelligence (AI) based predictive models using the data of pregnant patients as clinical parameters in the literature.

In this study, we aimed to determine the best-performing ML-based algorithm for predicting GDM with sociodemographic and obstetric features in the pre-conceptional period.

### Material-methods

We collected the maternal and pregnancy-related data of the patients who were admitted to the Eskisehir City Hospital, Obstetric Clinic between 01.01.2022 and 31.12.2023 retrospectively after ethical approval had been obtained. We identified the inclusion criteria as:

- Having hospital admittance in the first trimester of pregnancy.
- Being >18 and < 40 years old.
- Having spontaneous, non-anomalous, and singleton pregnancy.
- Having all data for the socio-demographic information of the pre-conceptional period (age, Body Mass Index (BMI), smoking, venous plasma glucose level, family history of diabetes mellitus), obstetrics history (gravida, parity, the previous birth weight), and the results of an oral glucose tolerance test.
- Not having a history of diabetes mellitus-related pregnancy disorders in previous pregnancies.
- Not having a history of a macrosomic delivery.
- Not having a diagnosis of diabetes mellitus.
- Being nulliparous or primiparous.

The maternal age, Body Mass Index (BMI), gravida, parity, previous birth weight, smoking status, the first-visit venous plasma glucose level, and the family history of diabetes mellitus were collected from the electronic hospital records retrospectively.

After determining the participants of the study with the inclusion criteria, the patients were classified as no GDM and GDM. The diagnosis of GDM was made if there was at least one abnormal value ( $\geq 92$ , 180, and 153 mg/dl for fasting, one-hour, and two-hour plasma glucose concentration, respectively, after a 75 g oral glucose tolerance test (OGTT) after the 24th gestational week. All

pregnancies included in the study were then grouped as nulliparous or primiparous.

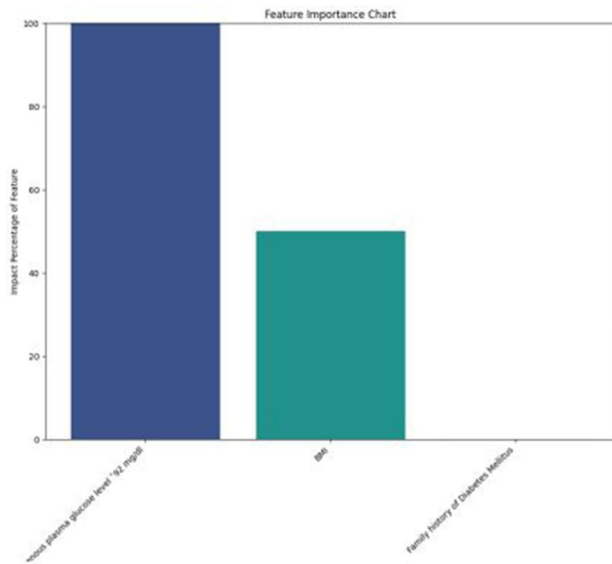
The Shapiro-Wilk test was used to test the normality of the variables in the analysis. Continuous variables were expressed as mean  $\pm$  standard deviation, and categorical variables as frequency (percentage) and median (minimum-maximum). The comparison of continuous variables that showed normal distribution between groups was made using the independent samples t-test, and the comparison of continuous variables that did not show normal distribution between groups was made using the Mann-Whitney U test. The relationship between categorical variables between groups was made using the Chi-Square test. In all analyses, a two-sided p-value  $< 0.05$  was considered statistically significant.

The study was designed according to the principles of ML. The Extra Trees Classifier, Average (AVG) Blender, Light Gradient Boosting Machine (LGBM) Classifier, eXtreme Gradient Boosting (XGB) Classifier, Logistic Regression, and Random Forest Classifier techniques were used as ML algorithms. For machine learning, 80% of the data (belonging to 80 pregnant women) was used for training and the remaining 20% was used for testing. In the tests conducted with these models, the model success rates were determined based on accuracy, sensitivity, and specificity values with confusion matrix metrics and the area under curve (AUC) graph in the receiver operating characteristic (ROC) curve analysis. Cross-validation is a statistical method used to estimate the ability of machine learning models and is widely used in applied machine learning to compare and select a model for a given predictive modeling problem. A confusion matrix contains information on actual and predicted classifications performed by a classification system and the performance of such systems is generally assessed using the data in the matrix. Independent variables that significantly affect the GDM dependent variable were selected by the permutation feature importance method, which is based on a decrease in the model's score when a single variable value is randomly shuffled. The Permutation Feature Importance Plot is given in Figs. 1, 2, 3 and 4.

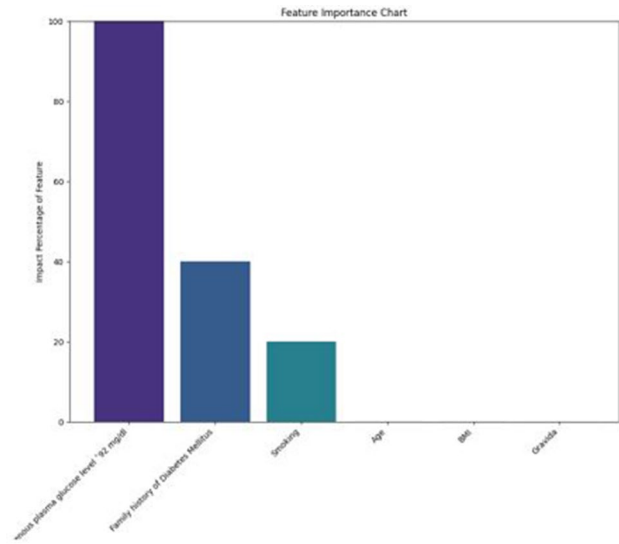
### Results

97 mothers were selected with strict criteria and included in the study (Fig. 5). These patients were classified as GDM detected and not detected according to the OGTT results. All mothers were then classified as nulliparous or primiparous. As shown in Tables 1, 19 and 26 nulliparous were with and without GDM, respectively whereas 29 and 23 primiparous were with and without GDM, respectively.

In the first part of the study, the nulliparous with and without GDM were analyzed statistically. The first-visit venous plasma glucose level and maternal BMI



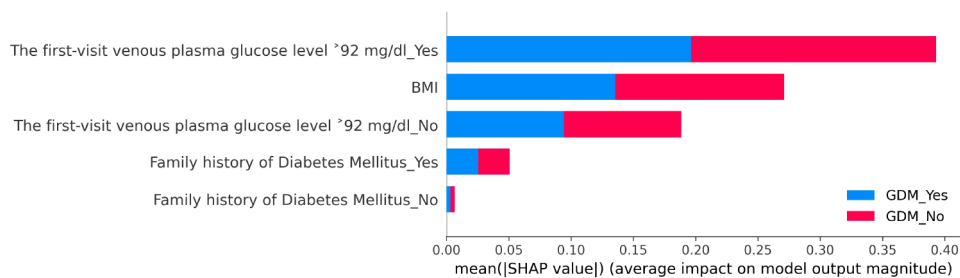
**Fig. 1** Feature importance charts of the group of the nulliparous' variables. 1: The first-visit fasting blood glucose level >92, 2: Body Mass Index, 3: the family history of diabetes mellitus



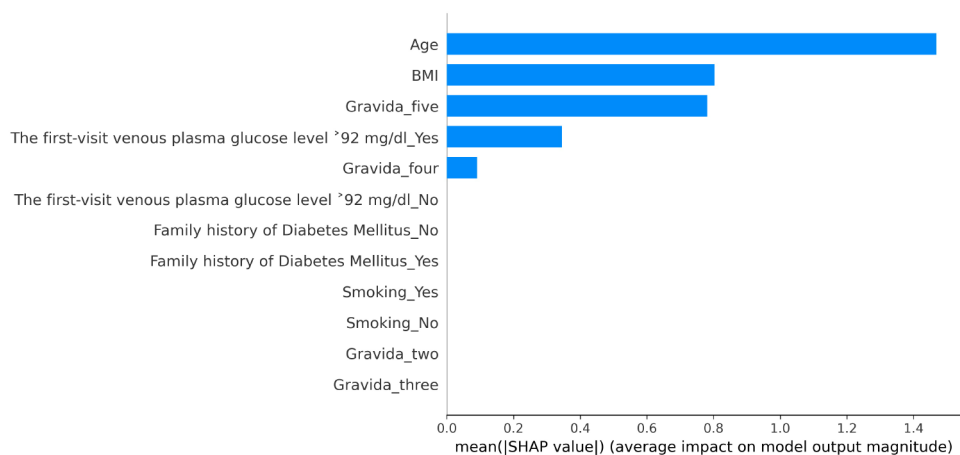
**Fig. 3** Feature importance charts of the group of the primiparous' variables. 1: The first-visit fasting blood glucose level >92, 2: the family history of diabetes mellitus, 3: Smoking, 4: previous birth weight > 4000 g, 5: Gravida, 6: Age, 7: Body Mass Index

had statistically significant differences between the two groups ( $p=0.017$  and  $0.002$ , respectively) as shown in Table 1, and these variables were the highest feature importance parameters among the six (Figs. 1 and 2).

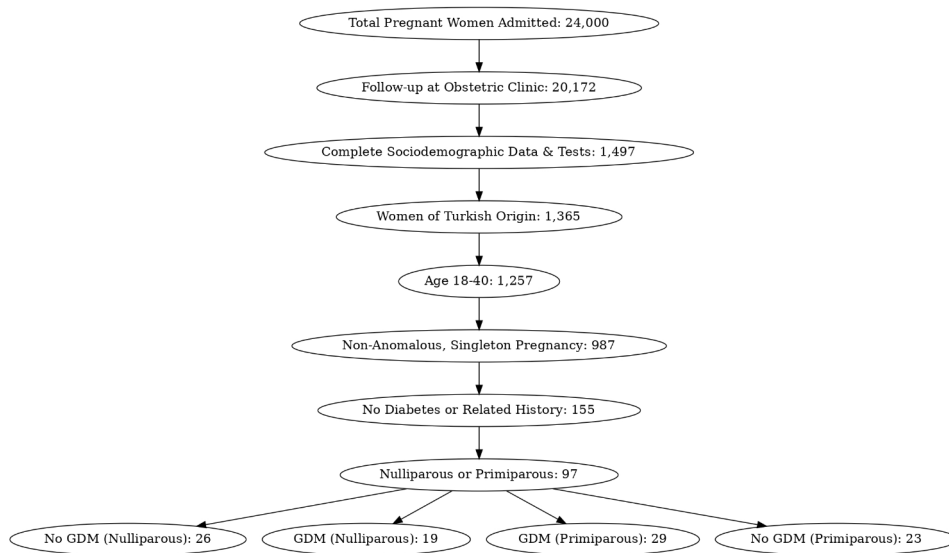
The ML algorithms were applied to the test data after the training dataset and it was found that the XGB Classifier model was the most powerful algorithm with the highest accuracy rate (66.7%) and AUC-ROC (55%) for predicting



**Fig. 2** SHapley additive explanations (SHAP) summary plot of feature selection model



**Fig. 4** SHapley additive explanations (SHAP) summary plot of feature selection model



**Fig. 5** Flow chart of the study

**Table 1** The demographic and obstetric findings of the patients included in the study

	Total Preg- nant Women	Nulliparous		p	Primiparous		p
		+	-		+	-	
GDM (n, %)	49.48%	19, 42.2%	26, 57.8%		29, 55.7%	23, 44.3%	
Age (years)	28.25 ± 5.08	27.16 ± 5.07	27.03 ± 5.29	0.940 <sup>a</sup>	29.21 ± 5.10	28.52 ± 4.45	0.614 <sup>a</sup>
Gravida (mean ± std)	2.06 ± 1.10	1.21 ± 0.41	1.11 ± 0.32	0.390 <sup>b</sup>	3.00 ± 1.07	2.52 ± 0.59	0.116 <sup>b</sup>
BMI (mean ± std)	30.77 ± 4.02	31.39 ± 3.90	28.65 ± 4.08	<b>0.017<sup>b</sup></b>	31.71 ± 3.70	31.35 ± 4.42	0.754 <sup>a</sup>
Smoking (n, %)				0.741 <sup>c</sup>			0.937 <sup>c</sup>
	+ 11 (11.3%)	2 (10.5%)	2 (7.7%)		4 (13.8%)	3 (13.0%)	
	- 86 (88.7%)	17 (89.5%)	24 (92.3%)		25 (86.2%)	20 (87.0%)	
Family History of Diabetes (n, %)				0.097 <sup>c</sup>			0.123 <sup>c</sup>
	+ 31 (32%)	6 (31.6%)	3 (11.5%)		15 (51.7%)	7 (30.4%)	
	- 66 (68%)	13 (68.4%)	23 (88.5%)		14 (48.3%)	16 (69.6%)	
The first-visit venous plasma glucose level >92 mg/dl				<b>0.002<sup>c</sup></b>			<b>0.010<sup>c</sup></b>
	+ 35 (36.1%)	9 (47.4%)	2 (7.7%)		18 (62.1%)	6 (26.1%)	
	- 64 (63.9%)	10 (52.6%)	26 (92.3%)		11 (37.9%)	17 (73.9%)	

GDM: Gestational diabetes mellitus, BMI: Body Mass Index

<sup>a</sup>Student t-test, <sup>b</sup>Mann-Whitney U (Data is not normal), <sup>c</sup>Chi-square

**Table 2** Prognosis prediction results of different machine learning algorithms

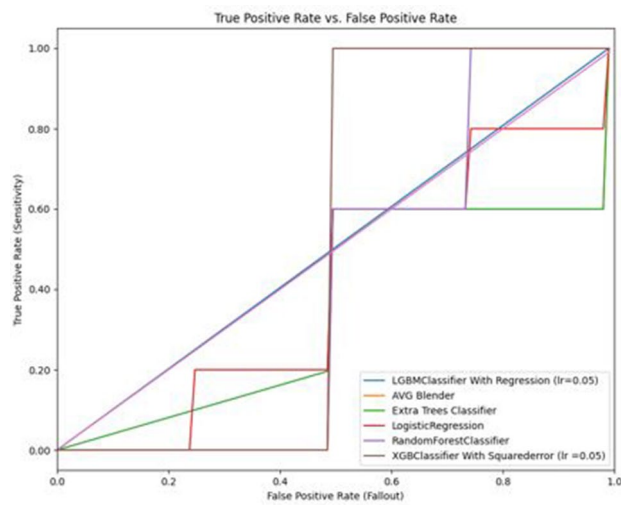
Model Name	Nulliparous				Primiparous			
	Model Training Results				Model Training Results			
	Test		Cross Validation		Test		Cross Validation	
Random Forest	ROC-AUC	Accuracy	ROC-AUC	Accuracy	ROC-AUC	Accuracy	ROC-AUC	Accuracy
Logistic Regression	0.525	0.667	0.686	0.639	0.55	0.636	0.587	0.571
LGBM Classifier	0.35	0.444	0.771	0.667	0.333	0.545	0.66	0.619
<b>XGB Classifier</b>	0.5	0.556	0.5	0.583	0.5	0.545	0.5	0.571
AVG Blender	<b>0.55</b>	<b>0.667</b>	0.629	0.639	<b>0.733</b>	<b>0.727</b>	0.59	0.619
Extra Trees Classifier	0.35	0.333	0.648	0.694	0.5	0.636	0.66	0.643
	0.425	0.667	0.762	0.639	0.933	0.818	0.562	0.5

AVG Blender: Average Blender, LGBM: Light Gradient Boosting Machine, XGB: eXtreme Gradient Boosting

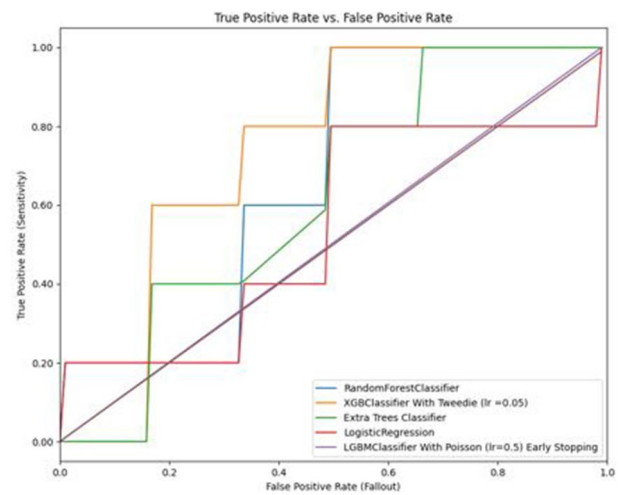
GDM using the maternal sociodemographic findings and the obstetric history (Table 2; Fig. 6), with sensitivity and specificity of 80 and 50%, respectively (Table 3).

In the second part of the study, the primiparous with and without GDM were analyzed statistically. Only the

first-visit venous plasma glucose level had statistically significant differences between the two groups ( $p=0.01$ ) as seen in Table 1, and the first-visit venous plasma glucose level >92 mg/dl, the family history of diabetes mellitus, and smoking were the highest feature importance



**Fig. 6** Receiver operating characteristic curve graphs of the nulliparous variables



**Fig. 7** Receiver operating characteristic curve graphs of the primiparous variables

parameters among the six (Figs. 3 and 4). The ML algorithms were applied to the test data after the training dataset and it was found that the XGB Classifier model was the most powerful algorithm with the highest accuracy rate (72.7%) and AUC-ROC (73.3%) for predicting GDM using the maternal sociodemographic findings and the obstetric history (Table 2; Fig. 7), with sensitivity and specificity of 40 and 100%, respectively (Table 3).

**Discussion**

GDM is one of the most common metabolic disorders of pregnancy and is related to the poor health outcomes of the pregnant woman/mother and/or the fetus/child. A pregnant woman with GDM would be more likely to have hypertensive disorders, cholestasis, and/or obstructed vaginal delivery. A mother with GDM has a higher chance of postpartum depression, kidney disease, type 2 diabetes, and/or cardiovascular problems in later life. While macrosomia, preterm delivery, and stillbirth could affect the fetus of a pregnant woman with GDM, neonatal hypoglycemia and diabetes are more likely to develop later in the life of a child of a pregnant woman with GDM [1].

Despite the onset of GDM being usual in the second trimester of the pregnancy and the OGTT being recommended during this period, early identification of a

high risk of developing the disease is important. Lifestyle changes, dietary regulation, and/or medication could be recommended to this group of pregnant women to reduce the incidence or severity of GDM-related disorders in the early and later future. The risk factors of GDM have been well-described in the literature [1], and in our study, the family history of diabetes, the venous plasma glucose level, the higher maternal age, and BMI were identified as risk factors. Risk predictive models and nomograms for GDM in the pre-conceptional period and/or first trimester have been constructed with sociodemographic characteristics and biochemical findings [4, 10–12]. However, the main limitations of these early screening programs were that the sample sizes were small, most did not have a prospective multi-center design, and they were not externally validated. To our knowledge, the only externally validated model for predicting GDM was The Monash model [4, 6]. A systematic review and meta-analysis of the screening programs declared that pre-diagnostic predictive models constructed with one, two, or several risk factors in the literature had poor performance and a lack of evidence [13].

ML algorithms are viewed as a possible solution to construct predictive models with high performance in order to overcome the above-mentioned limitations. It is well known that ML algorithms have more advantages than

**Table 3** Classifier confusion matrix

	XGB Classifier			Primiparous		
	Yes	No	%	Yes	No	%
Yes	4	1	80.0 <sup>a</sup>	2	3	40.0 <sup>a</sup>
No	2	2	50.0 <sup>b</sup>	0	6	100.0 <sup>b</sup>

XGB: eXtreme Gradient Boosting, GDM: Gestational diabetes mellitus

a: sensitivity b: specificity



traditional statistical methods as they can uncover relationships independent from the data structures and can provide highly stable predictions in small or large datasets where correlations are often difficult to find. In addition, researchers have various ML models to help find the optimal algorithm [14, 15]. The ML-based models for predicting GDM in the first trimester were reviewed previously [4]. The sociodemographic characteristics, biochemical and ultrasonographic parameters, the family history of diabetes, and the obstetric findings were identified as variables to construct the models. Ye et al. did not detect any statistically significant differences in the prediction performance of ML algorithms for developing GDM, for this reason, the researchers chose conventional logistic regression methods [16]. No independent external validated studies for ML-based predictive models have been reported, and more reports are still being published to identify ML-based models with high accuracy and AUC [17–19].

ML-based predictive models using women's health status in the preconceptional period have more importance especially for women in remote or low-income rural areas, because, as is well known, it could be difficult for pregnant women living in these areas to be admitted to hospital. As a result, the risk classification of developing GDM can be based only on the patient's medical, obstetric, and family history instead of any biochemical or ultrasonographic parameters. It was also recommended that clinical needs and achievable parameters are more important than the accuracy of clinical decision support systems; as a result, excessive feature parameters that are difficult to obtain in routine medicine should not be selected for model construction [20]. For these reasons, we tried to construct predictive models based on ML with basic medical, obstetric, anthropometric, and family history variables. To create an effective model, we evaluated pregnant women with or without a history of live birth in separate groups. In our study, the greatest feature importance variables were the first-visit blood glucose level, maternal BMI, and the family history of diabetes mellitus. These variables were also identified as significant risk factors for GDM [21] and have been selected as the strongest feature importance factors for developing ML models previously [20]. The XGB Classifier with 75.7% accuracy and 74.2% ROC-AUC [17], Artificial Neural Network with 70.3% accuracy and 83.3% sensitivity [18], Logistic Regression with 72.8% ROC-AUC and 64.9% sensitivity [19], and Random Forest with 79.9% ROC-AUC and 75% sensitivity [22] have been detected previously as the best-performing algorithms for predicting GDM with the same feature important variables as our study in the preconceptional or early period of pregnancy. The most common ML algorithms (The Extra Trees Classifier, AVG Blender, LGBM Classifier, XGB

Classifier, Logistic Regression, and Random Forest Classifier) were employed to develop the prediction model and these achieved a range of accuracy and ROC-AUC of 55–70% and 45–73%, respectively. Among these algorithms, we found that the XGB Classifier had the best predictive value for the two models (Table 2). The accuracy of 77.8% and the 80% sensitivity indicate that our model is highly effective in identifying all true positive nulliparous instances of GDM (i.e., all women with GDM were correctly identified by the model). This high sensitivity is critical in a healthcare context where failing to identify a GDM case could have significant negative consequences for both the mother and the baby. However, the moderate AUC value of 0.55 suggests that the model may not be distinguishing well between true positive and false positive cases across all thresholds. This could be due to the nature of our data and the balance between GDM and non-GDM cases. In particular, the AUC score can be sensitive to the distribution of the data, especially in cases where there is an imbalance or specific characteristics in the feature space that make differentiation challenging. The high sensitivity and accuracy indicate that while the model effectively captures true positive GDM cases, it may be less effective at avoiding false positives. This outcome aligns with our focus on ensuring that no GDM cases are missed (high sensitivity) even if it comes at the cost of accepting more false positives. We believe that this trade-off is acceptable and even desirable in the context of GDM prediction, where the priority is to ensure that all at-risk pregnancies are identified for further monitoring and intervention. However, we acknowledge that there is room for improvement in fine-tuning the model to better balance sensitivity and specificity, which could improve the AUC score. The low specificity of 50% suggests that the model identifies a substantial number of false positive cases where the model predicts GDM, but the individual does not actually have the condition. While this might lead to some women receiving additional monitoring or tests unnecessarily, the potential benefits of catching every possible GDM case outweigh the drawbacks in this context. This trade-off is often considered acceptable in screening scenarios, where the goal is to minimize the risk of missing any true cases. In our study, this balance between sensitivity and specificity reflects our prioritization of ensuring that all GDM cases are captured. This approach is particularly relevant in populations where the cost of missed diagnoses is high. We recognize that low specificity may lead to a higher number of false positives, and we aim to address this in future model refinements by exploring ways to improve specificity without compromising sensitivity.

The limitations of this study were that the dataset was limited and consisted of only Turkish pregnant patients. It is important to note that while small sample sizes can

indeed pose challenges in traditional statistical methods, this is not necessarily the case in machine learning. Machine learning models, especially when properly tuned and validated, are capable of extracting meaningful patterns from limited data, provided that the data is of high quality and the feature set is appropriately selected. In our study, we employed techniques to mitigate the risks associated with a small dataset, such as cross-validation, which helps ensure that the model's performance is not overly optimistic and that it generalizes unseen data well. Furthermore, our model demonstrated robust performance metrics, which we believe validate its reliability even with the smaller sample size. Another point is that 48 out of the 97 women (49.48%) were diagnosed with GDM. This high prevalence was observed in a cohort even after excluding women with previous GDM and those who gave birth to macrosomic babies. This prevalence may appear elevated, particularly when compared to general population statistics. However, we intentionally included an equal number of women with and without GDM to ensure a balanced comparison between the two groups. Our primary aim was to clearly identify the relationships between various parameters in women with GDM, which might have been less discernible with a more unbalanced cohort. This prevalence might also suggest a selection bias or limit the generalizability of our findings to the broader Turkish population, however, the study was not intended to reflect the actual prevalence of GDM in the general population. Instead, it was structured to optimize the training of the machine learning model by providing an equal representation of GDM and non-GDM cases. This approach was necessary to enhance the model's ability to distinguish between the two groups effectively. In addition, we excluded women with previous GDM and a history of delivering a macrosomic baby in a previous pregnancy. The rationale behind these exclusions was to focus our analysis on a population of pregnant women who had no prior history of GDM or related complications, thus minimizing potential confounding factors. By excluding women with known risk factors or previous diagnoses, we aimed to create a cohort where the development of GDM in the current pregnancy could be more closely associated with the parameters being studied, rather than with pre-existing conditions. This approach allowed us to identify potential indicators of GDM in a population that would not typically be flagged as high-risk based on their obstetric history. Furthermore, including women with a history of GDM or macrosomia could have affected the results, as these women are already more likely to develop GDM in subsequent pregnancies. Our goal was to examine the development of GDM in a population without these known predispositions, thereby highlighting the

predictive power of the model in a broader and more generalizable group of women.

Other variables, such as lifestyle behaviors, which may play a role in the development of GDM, could not be included because of the retrospective design of the study. In addition, external validation of the dataset was not used to evaluate the predictive model. We identified having a history of diabetes mellitus-related pregnancy disorders and macrosomic baby delivery as exclusion criteria because we could not obtain the data about these subjects from the electronic medical records available. The strength of this study is that we showed that the basic socio-demographic characteristics, the family history of diabetes, and only one venous plasma glucose level could have sufficient capacity to develop a predictive model for GDM. A further strength of this study is that the model we developed is specifically tailored to the Turkish population, which helps to minimize potential confounding factors related to ethnic diversity. By focusing on a homogeneous population, we were able to eliminate the complexities and variations that may arise when applying a generalized model across different ethnic groups. This approach enhances the model's accuracy and reliability within the Turkish context, making it more relevant and applicable for GDM prediction in this specific setting. Moreover, many existing machine-learning models for GDM prediction have been developed primarily for Caucasian and Asian populations [23]. The value of our study lies in its ability to address the unique demographic and genetic factors present in the Turkish population, which may differ from those in other ethnic groups. By highlighting this advantage, this research contributes to the growing body of research that recognizes the importance of developing population-specific models in the field of personalized medicine.

The interest in AI-based clinical decision support systems is increasing all over the world and these systems represent a major opportunity to improve the health of people living in countries with low levels of income. Externally validated predictive models based on ML algorithms and with large datasets constructed from different ethnicities with variables related to medical-obstetric family history, lifestyle factors, and physical information would be useful for the screening, diagnosis, and risk assessment of GDM in the pre-conceptual period or the first trimester and could improve maternal and fetus/child health. This work serves as an initial step toward developing a reliable model for the early prediction of gestational diabetes, with the potential for refinement and validation in subsequent studies involving larger populations. In addition, there is a future research possibility of developing a similar machine learning model for GDM to postpartum Type 2 Diabetes progression for more effective early interventions in public health

settings. The transition from GDM to postpartum Type 2 diabetes is a critical public health issue, as women with a history of GDM are at an increased risk of developing Type 2 diabetes later in life. Developing a machine learning model that predicts this progression could significantly enhance early intervention strategies and improve long-term health outcomes for these women.

#### Author contributions

YK and ZB designed the study. ES, YK and TT processed the data and wrote the manuscript. AY and ÖÇ performed data analysis. ZB xxxxxx critically reviewed the final version of the manuscript. All authors contributed to the article and approved the submitted version.

#### Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

#### Data availability

The data underlying this article will be shared upon reasonable request to the corresponding author.

#### Declarations

#### Ethics approval

The study was approved by the Ethical Committee of the Ankara Medipol University (12.02.2024/ 17).

#### Consent to participate

Informed consent was not obtained from participants due to the retrospective study design.

#### Competing interests

The authors declare no competing interests.

Received: 28 June 2024 / Accepted: 26 August 2024

Published online: 31 August 2024

#### References

- McIntyre HD, Catalano P, Zhang C, Desoye G, Mathiesen ER, Damm P. Gestational diabetes mellitus. *Nat Reviews Disease Primers*. 2019;5(1):47.
- Sohmaran C, Bte Mohamed Rahim A, Chua JYX, Shorey S. Perceptions of primiparous women diagnosed with gestational diabetes mellitus: a descriptive qualitative study. *Midwifery*. 2013;125:103802. <https://doi.org/10.1016/j.midw.2023.103802>.
- Sweeting AN, Ross GP, Hyett J, Wong J. Gestational diabetes in the first trimester: is early testing justified? *Lancet Diabetes Endocrinol*. 2017;5(8):571–3. [https://doi.org/10.1016/S2213-8587\(17\)30066-9](https://doi.org/10.1016/S2213-8587(17)30066-9).
- Cooray SD, De Silva K, Enticott JC, Dawadi S, Boyle JA, Soldatos G, Paul E, Versace VL, Teede HJ. Temporal validation and updating of a prediction model for the diagnosis of gestational diabetes mellitus. *J Clin Epidemiol*. 2023;164:54–64. <https://doi.org/10.1016/j.jclinepi.2023.08.020>.
- Lamain e de Ruiter M, Kwee A, Naaktgeboren CA, Franx A, Moons KGM, Koster MPH. Prediction models for the risk of gestational diabetes: a systematic review. *Diagn Progn Res*. 2017;1:3.
- Teede HJ, Harrison CL, Teh WT, Paul E, Allan CA. Gestational diabetes: development of an early risk prediction tool to facilitate opportunities for prevention. *Aust N Z J Obstet Gynaecol*. 2011;51:499e504.
- Kumar M, Chen L, Tan K, Ang LT, Ho C, Wong G, Soh SE, Tan KH, Chan JKY, Godfrey KM, Chan SY, Chong MFF, Connolly JE, Chong YS, Eriksson JG, Feng M, Karnani N. Population-centric risk prediction modeling for gestational diabetes mellitus: a machine learning approach. *Diabetes Res Clin Pract*. 2022;185:109237. <https://doi.org/10.1016/j.diabres.2022.109237>.
- Artzi NS, Shilo S, Hadar E, Rossman H, Barbash-Hazan S, Ben-Haroush A, et al. Prediction of gestational diabetes based on nationwide electronic health records. *Nat Med*. 2020;26(1):71–6.
- Wu Y-T, Zhang C-J, 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 Metabol*. 2020;106:e1191–205.
- Gerszi D, Orosz G, Török M, Szalay B, Karvaly G, Orosz L, Várbiro S. Risk estimation of gestational diabetes mellitus in the first trimester. *J Clin Endocrinol Metabolism*. 2023;108(11):e1214–23.
- Wang X, He C, Wu N, Tian Y, An S, Chen W, Shen X. Establishment and validation of a prediction model for gestational diabetes. *Diabetes Obes Metabolism*. 2024;26(2):663–72.
- Zhang H, Dai J, Zhang W, Sun X, Sun Y, Wang L, Li H, Zhang J. Integration of clinical demographics and routine laboratory analysis parameters for early prediction of gestational diabetes mellitus in the Chinese population. *Front Endocrinol (Lausanne)*. 2023;13(14):1216832. <https://doi.org/10.3389/fendo.2023.1216832>.
- Farrar D, Simmonds M, Bryant M, Lawlor DA, Dunne F, Tuffnell D, Sheldon TA. (2017). Risk factor screening to identify women requiring oral glucose tolerance testing to diagnose gestational diabetes: A systematic review and meta-analysis and analysis of two pregnancy cohorts. *PLoS One*. 6;12(4):e0175288. <https://doi.org/10.1371/journal.pone.0175288>
- Byambasuren O, Beller E, Glasziou P. Current knowledge and adoption of mobile health apps among Australian general practitioners: survey study. *J JMIR mHealth uHealth*. 2019;7(6):e13199.
- Shukla VV, Eggleston B, Ambalavanan N, et al. Predictive modeling for perinatal mortality in resource-limited settings. *JAMA Netw Open*. 2020;3(11):e2026750.
- Ye Y, Xiong Y, Zhou Q, Wu J, Li X, Xiao X. Comparison of machine learning methods and conventional logistic regressions for Predicting Gestational Diabetes using Routine Clinical Data: a retrospective cohort study. *J Diabetes Res*. 2020;12:2020:4168340. <https://doi.org/10.1155/2020/4168340>.
- Liu H, Li J, Leng J, Wang H, Liu J, Li W, Liu H, Wang S, Ma J, Chan JC, Yu Z, Hu G, Li C, Yang X. Machine learning risk score for prediction of gestational diabetes in early pregnancy in Tianjin, China. *Diabetes Metab Res Rev*. 2021;37(5):e3397. <https://doi.org/10.1002/dmrr.3397>.
- Gallardo-Rincón H, Ríos-Blancas MJ, Ortega-Montiel J, Montoya A, Martínez-Juarez LA, Lomelín-Gascón J, Saucedo-Martínez R, Mújica-Rosales R, Galicia-Hernández V, Morales-Juárez L, Illescas-Correa LM, Ruiz-Cabrera IL, Díaz-Martínez DA, Magos-Vázquez FJ, Ávila EOV, Benítez-Herrera AE, Reyes-Gómez D, Carmona-Ramos MC, Hernández-González L, Romero-Islas O, Muñoz ER, Tapia-Conyer R. MIDO GDM: an innovative artificial intelligence-based prediction model for the development of gestational diabetes in Mexican women. *Sci Rep*. 2023;28(131):6992. <https://doi.org/10.1038/s41598-023-34126-7>.
- Lee SM, Hwangbo S, Norwitz ER, Koo JN, Oh IH, Choi ES, Jung YM, Kim SM, Kim BJ, Kim SY, Kim GM, Kim W, Joo SK, Shin S, Park CW, Park T, Park JS. Non-alcoholic fatty liver disease and early prediction of gestational diabetes mellitus using machine learning methods. *Clin Mol Hepatol*. 2022;28(1):105–16. <https://doi.org/10.3350/cmh.2021.0174>.
- Zhang Z, Yang L, Han W, Wu Y, Zhang L, Gao C, ... & Wu H. (2022). Machine learning prediction models for gestational diabetes mellitus: meta-analysis. *Journal of medical Internet research*, 24(3), e26634.
- Sweeting A.N., Appelblom H., Ross G.P., Wong J., Kouru H., Williams P.F., ... Hyett J.A. (2017). First trimester prediction of gestational diabetes mellitus: a clinical model based on maternal demographic parameters. *Diabetes research and clinical practice*, 127, 44–50.
- Huang Z, Ruppenkamp J, Krishnan D, Holy CE. AI3 discriminative ability of commonly used indices to predict outcomes after total knee replacement: a comparison of demographics, provider volume, ASA score, Charlson, elixhauser and functional comorbidity index. *Value Health*. 2019;22:S34.
- Kumar M, Ang LT, Ho C, Soh SE, Tan KH, Chan JKY, Godfrey KM, Chan SY, Chong YS, Eriksson JG, Feng M, Karnani N. Machine learning-derived prenatal predictive risk model to Guide Intervention and Prevent the Progression of Gestational Diabetes Mellitus to Type 2 diabetes: Prediction Model Development Study. *JMIR Diabetes*. 2022;7(3):e32366.

#### Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.