Journal of Artificial Intelligence in Medical Applications (JAIMA)
Vol. 1, No. 1, 2025, **pp.** 55–63
ISSN: 3106-1494 | **DOI**: https://doi.org/10.65511/jaima.v1i1.739

A Machine Learning Technique for Early Detection of Gestational Diabetes Mellitus Using SMOTE and Optimized Light Gradient Boosting Machine

Ahmed Adil Nafea[1, *], Mohammed M AL-Ani[2], Afrig Aminuddin[3], Meaad Ali Khalaf[4], Amani Steiti[5], Aythem Khairi Kareem[6], Saeed Amer Alameri[7]

[1]Department of Artificial Intelligence, College of Computer Science and IT, University of Anbar, Ramadi, Iraq; ahmed.a.n@uoanbar.edu.iq
[2]Center for Artificial Intelligence Technology (CAIT), Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia (UKM), Bangi, Selangor, Malaysia; mohmed.alanni@yahoo.com
[3]Faculty of Computer Science, Universitas Amikom Yogyakarta Sleman, 55283, Indonesia. afrig@amikom.ac.id
[4]Department of Computer Science, AUL University, Beirut, Lebanon; meaad.ali@gmail.com
[5]Department of Computer Systems and Networks, Faculty of Information Engineering, University Tishreen, Latakia, Syria; amanystiety1@gmail.com
[6]Department of Heet Education, General Directorate of Education in Anbar, Ministry of Education, Heet, 31007 Anbar, Iraq; ayt19c1004@uoanbar.edu.iq
[7]College of Computers, Seiyun University, Hadhramout, Yemen; salameri@seiyunu.edu.ye

Received: 13/06/2025, Revised: 11/08/2025, Accepted: 24/08/2025, Published: 16/09/2025

ABSTRACT: Gestational Diabetes Mellitus poses important health risks to both mother and child if not detected and managed early. The timely detection is important for minimizing outcomes ranging from preeclampsia and high baby birth weight to later onset of type 2 diabetes. This manuscript outlines a robust and interpretable machine learning pipeline to predict GDM, utilizing clinical data acquired from pregnant women from Kurdistan region of Iraq. The SMOTE technique is incorporated for tackling the imbalanced class problem over the dataset and to improve the model performance to detect minority class probes of GDM. We also use the Light Gradient Boosting Machine (LightGBM) as our classifier because of its fast speed and high accuracy; and the hyper-parameters are searched by automatically by Optuna, a widely used hyper-parameter optimization framework. The model was tested in two different ratios of train-test split (70%–30% and 80%–20%), and it achieved high accuracy consistently (the highest model performance accuracy reached up to 91%). These results demonstrate the advanced predictive power of the optimized model over the baseline and pragmatic application approaches. Practicality of this approach will lead to early detection of GDM timely interventions for pregnancy outcome thereby encouraging an approach against mother and fetus health.

Keywords: Gestational Diabetes Mellitus, Machine Learning, LightGBM, Hyperparameter Optimization, Medical Diagnosis, Early Detection.

1. INTRODUCTION

Gestational Diabetes Mellitus (GDM) is a type of diabetes that happens through pregnancy [1]. Its defined as an elevated blood glucose level during gestation that does not meet all diabetes diagnostic criteria prior to gestation [2]. It is a common non-communicable obstetrical complexity in pregnancy. The majority of GDM is growing quickly worldwide. In the year of 2021, the moderate majority of GDM was reported to be 16.7% globally, based on geographical locations, diagnostic criteria, and population [3]. GDM is related to both long- and short-term adverse health effects for children and mothers. During gestation, GDM is related to an enhanced hazard of pre-eclampsia and heightened fetal development, leading to neonatal hypoglycemia, birth trauma, shoulder dystocia, and macrosomia. Women with GDM also have a 50% hazard of developing type two diabetes in the decade following their gestation [4]. Use a 2-hour fasting 75-g criteria for diagnosis of GDM at 24-28 weeks gestation. Use the International Association of Diabetes and Pregnancy Study Groups (IADPSG) criteria or American Diabetes Association (ADA) criteria (IADPSG is preferred) to diagnose GDM at 24-28 weeks gestation. Previous research shows that the diagnosis of GDM at 24–28 weeks' gestation may be too late for intervention. Consequently, earlier diagnosis of GDM, as well as appropriate control, might be important, especially for overweigh and aged mothers with high hazard of GDM in order to control fetal abdominal obesity [5].

Accurate detection of GDM risk is important for early intervention and detection to avoid adverse outcomes. Nevertheless, predicting GDM risk might be challenging on account of the complex relationship of different risk factors such as prior history of GDM, family history of diabetes, maternal age, and ethnicity. Further, the precision of GDM risk predictions is restricted with the exposition time-varying characteristics that were estimated from risk factors measured at only one time point; these fail to consider the dynamic variations in clinical measures as seen during pregnancy. The difficulty in recommending personalized recommendations of medicine for GDM arises from the heterogeneity of the disease and absence of a consensus for the best treatment practice. New management approaches for GDM include different medications, activity proposals and dietary regimens [6]. Nonetheless, the current diagnostic methods are expensive and may overlook early symptoms. Current predictive methods have commonly had limited clinical features and poor performance because of limited generalizability and poor interpretability.

To address these challenges, machine learning (ML) has become a game-changing paradigm in healthcare, which has the potential to consider the complex interplay of various clinical parameters. Numerous AL algorithms, including LR, DT, SVM, RF and ensemble learning show promise in GDM prediction [7],[8],[9]. Nonetheless, they suffer issues such as imbalance data, lack of hyperparameter tune and low interpretability. To address these gaps, this study proposed an optimized ML framework for the early detection of GDM. The framework incorporates two key enhancements: SMOTE for addressing class imbalance and Optuna, an automated hyperparameter optimization framework, to fine-tune the LightGBM classifier. LightGBM was selected for its computational efficiency, scalability, and robustness in handling high-dimensional clinical data. Therefore, the main contributions of our Paper can be listed as follows:

- Implementation of SMOTE for handling data imbalance, so that minority class predicted output gets better sensitivity of model and reduced bias in the classification.
- Introduction of the Optuna for automatic and intelligent hyperparameter tuning allowing for better generalization power and accuracy in the LightGBM classifier.
- LightGBM implementation, an efficient, scalable and fast Gradient Boosting Decision Tree framework designed for clinical datasets, and able to deal effectively with high-dimensional structured data.
- All-rounded model testing with proven performance metrics like accuracy, precision, recall, F1-score, and confusion matrix for both train-test split scenarios (70–30 and 80–20).

This paper is structured as follows: Section 2 shows the related work on GDM prediction by employing ML. The proposed methodology, comprising the data pre-processing, SMOTE balancing, and Optuna-based LightGBM optimization, is summarized in Section 3. Section 4 gives results and discusses comparisons with other methods. Section 5 summarizes the study and points to future research.

2. RELATED WORKS

Numerous studies have investigated the prediction of GDM using ML techniques. These studies have used different techniques and yielded varying results. This section discusses these studies. Saleh et al. [10] proposed a ML approach to prediction of GDM, with four classification algorithms applied, namely LR, DT, GB and SVM. The model was constructed over PIMA diabetes dataset consisting of nine clinical parameters for each patient. In order to address missing data, we used a hybrid method of median and averaged imputation, with which three preprocessed data were obtained. The classifier was evaluated in terms of the commonly used evaluation parameters such as precision, specificity, accuracy, and sensitivity. The reliability and accuracy of such an approach were validated using the experimental results.

Alapati et al. [11] proposed ML and deep learning (DL) methods to improve the prediction of GDM. Using a comprehensive dataset that included glucose levels, medical history, and demographic data, the study compared the performance of a DL model with traditional ML techniques such as DT and SVM. Model evaluation was conducted utilizing metrics like accuracy, F-score, recall, and precision. The findings suggest that DL has strong potential to enhance GDM prediction and supports the development of more effective screening methods to improve maternal and fetal health results. Garg et al. [12] proposed many ML techniques to predict GDM, assessing six techniques: XGBoost, DT, LR, SVM, KNN, and RF. The GDM dataset contains features such as age, skin thickness, insulin levels, blood pressure, diabetes pedigree function, pregnancies, glucose levels and label. Through exhaustive experiments, these techniques are evaluated based on the RMSE, R-squared, MSE, and accuracy, with the outcomes visually demonstrated to facilitate enhanced understanding. The proposed approach aims to enhance GDM management and prediction, providing more informed clinical decisions and accessible health monitoring tools while also enhancing healthcare products for pregnant women.

Zhao et al. [13] presented an integrated learning strategy to build a method that high accurately forecasts GDM. The presented approach used multiple ML techniques to create classifiers and integrated them into a more effective model using Stacking approaches. These base ML classifiers include CatBoost, XGBoost, SVM, and RF. The presented approach utilized real, publicly available GDM datasets and improved the model's performance through feature selection and rigorous data preprocessing. The results of this analysis demonstrate the superiority and effectiveness of the Stacking approach in predicting GDM, presenting a possible explanation for early prediction and intervention. The results of this analysis are of great importance to patients, clinical doctors, and public health workers. Bhuria et al. [14] developed and assessed an ML technique based on LR to predict GDM using a large dataset of physiological characteristics and health features. The presented approach aims to develop a binary classification model that utilizes features such as family history, age, insulin levels, blood pressure, and pregnancy status. The results shows the efficiency of the LR technique, achieving 77% accuracy.

Shetty et al. [15] proposed a practical analysis utilizing many ML methods to forecast the prospective hazard characteristics affecting GDM progression. The ML techniques that are used are LR, SVM, RF, Gaussian NB, XGBoost, Voting Ensemble, Bagging with DT Weighted Average Ensemble, and Stacking XGBoost, NGBoost, AdaBoost. Most of the presented classifiers reached a suitable accuracy range of (75%-82 %). Kaya et al. [16] collected data from expectant women who were accepted to the obstetric clinic in the first three trimesters. The dataset includes several features, such as diabetes family history, glucose level, smoking status, previous birth weight, parity, gravida, body mass index, maternal age, and results of an oral glucose tolerance test, which were assessed for the patients. The presented approach employs six standard ML techniques, including LGBM, XGBoost, RF, LR, Average (AVG) Blender, and Extra Trees (ET). The XGBoost had the best predictive significance, with an accuracy of 66.7%. Victor et al. [17] presented an intelligent approach based on ML techniques to predict GDM, including XGBoost, CatBoost, AdaBoost, LGBM, and RF. The presented approach investigated data from Brazil, which included 1,557 expectant women with GDM at 19 weeks or less. The RF technique yielded the best-performing model, achieving an accuracy of 76%. Bigdeli et al. [5] collected a dataset of pregnant women from (2020-2022). Risk characteristics for developing detection models were determined through the input of clinical specialists and expert opinions, as well as a literature review. The extracted information experienced preprocessing with six ML techniques were employed and evaluated for GDM forecasting, including DT, Multi-Layer Perceptron, KNN, NB, XGBoost and RF. The presented approach was evaluated using Area Under Curve, sensitivity, accuracy, and precision. The RF technique achieved excellent performance in GDM forecasting, with recall of 92%, accuracy of 89%, and precision of 86%. The finding of this analysis indicate that ML techniques, specifically RF, have sufficient accuracy in the detection GDM.

Existing literature highlights a wide range of ML techniques applied to GDM prediction, with varying levels of success. While several models achieved good accuracy, challenges like class imbalance, limited hyperparameter tuning, and lack of optimization techniques remain prevalent. To address these gaps, this study proposes an enhanced LightGBM-based model optimized via Optuna and balanced using SMOTE to improve the accuracy of detection.

3. RESEARCH METHODOLOGY

This study proposes a ML framework for the early prediction of GDM utilizing an optimized LightGBM model combined with data balancing and hyperparameter tuning techniques. As shown in Figure 1, the methodology follows a systematic pipeline that includes data preprocessing, handling class imbalance, feature scaling, model tuning using Optuna, and performance evaluation using appropriate metrics. The workflow illustrated in figure 1 outlines a machine learning pipeline designed for classification tasks involving imbalanced datasets. Initially, the raw data undergoes preprocessing using StandardScaler to normalize the features, followed by the application of the SMOTE technique to address class imbalance by synthetically generating minority class samples. The preprocessed and balanced data is then split into two branches—one used for model training and the other for validation. Hyperparameter tuning is carried out using the Optuna optimization framework, which explores a defined search space for parameters such as learning rate, number of estimators, max depth, and regularization terms to maximize model performance. After optimization, the best set of parameters is used to train the LightGBM classifier. Finally, the model's effectiveness is evaluated using key metrics such as precision, recall, F1-score, and accuracy to ensure robust performance across both majority and minority classes.

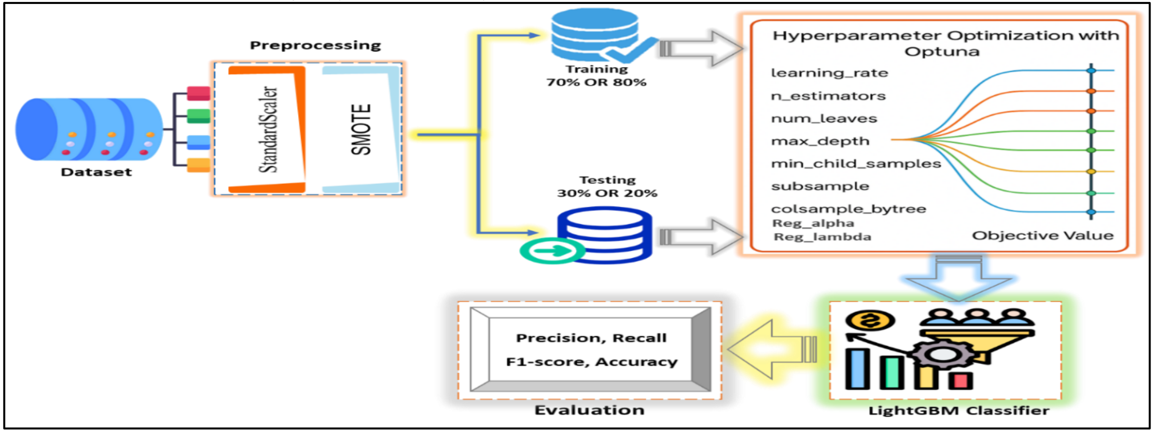

Figure 1. Proposed methodology.

3.1 Dataset Description

Comprises clinical and demographic data from pregnant women diagnosed with or without GDM [18]. The dataset consists of 1,012 records and includes seven variables, of which six are input features, and one is the target label. These features are Age, Number of Pregnancies, Weight, Height, Body Mass Index (BMI), and Heredity, which indicates whether the individual has a family history of diabetes. The target variable, labeled as Prediction, is binary, with 0 representing healthy individuals and 1 indicating a positive diagnosis of gestational diabetes.

3.2 Handling Class Imbalance with SMOTE

The dataset exhibited significant class imbalance, with 795 samples belonging to the non-diabetic class and only 217 samples representing gestational diabetes cases. This imbalance posed a risk of bias in ML model training, favoring the majority class. To mitigate this, the SMOTE was applied. SMOTE generated synthetic examples of the minority class to balance the distribution, resulting in 795 samples per class, as illustrated in Figure 2. This balancing step was important for enhancing model performance and ensuring fair learning across both classes.

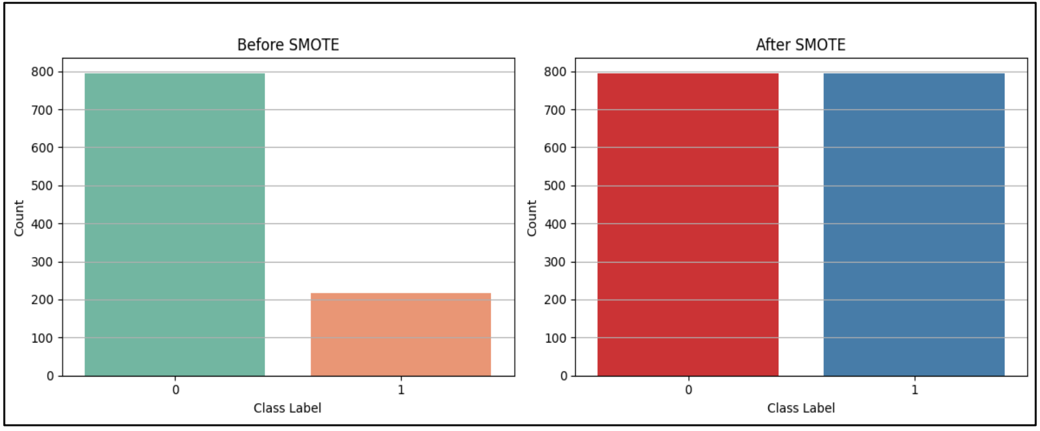

Figure 2. Class distribution before and after SMOTE

3.3 Feature Scaling

The dataset was splitting into testing and training data utilizing the ratio of (20%-80%) and (30%-70%). To normalize the input features and facilitate model convergence, in this work normalized the data to be with the mean of 0 and variance of 1 by StandardScaler [19], [20]. This step is very vital for gradient based models, like lightgbm to work well when the features are of varying magnitude.

3.4 LightGBM Optimization with Optuna

In order to improve the predictive ability of the developed ML model, an advanced hyperparameter optimization framework called Optuna was utilized in this study to optimize the parameters of the LightGBM. The LightGBM is a high-performance, distributed gradient boosting framework based on decision tree algorithms. It is designed for speed and efficiency, making it well-suited for tasks such as classification, ranking, and other machine learning applications. Hyperparameter optimization is crucial to achieve the maximum generalization with ML models. Suboptimal parameters set up underfitted or overfitted context, and both degrade performance. To overcome this, Optuna automates the optimal hyperparameter search with TPE, an effective sampling algorithm. The approach dynamically samples configurations to visit, and learns a model of the distribution of successful configurations, rather than simply choosing configurations uniformly at random or in a grid. The hyperparameter settings for lightGBM via Optuna optimization show in Table 1.

Table 1. Hyperparameter settings via Optuna optimization.

Parameter	Optimized Range (via Optuna)
Num_leaves	20 – 150
Min_child_samples	5 – 30
Subsample	0.5 – 1.0
Max_depth	3 – 12
Colsample_bytree	0.5 – 1.0
Learning_rate	0.01 – 0.2
N_estimators	100 – 300
Reg_lambda	0.0 – 5.0
Reg_alpha	0.0 – 5.0

Optimization is formulated as an objective function that takes a set of hyperparameters proposed by Optuna in a trial and computes its validation accuracy after training the LightGBM model on the training data. The goal was to increase accuracy as it is the simplest and most intuitive performance measure for binary-classification tasks. We run 30 trials in total, where each trial trains and tests a different LightGBM model configuration. Optuna would repeatedly change the hyperparameter settings and see how it would change the model's accuracy throughout each trial. By taking advantage of knowledge from previous experiments, Optuna effectively reduced the search space and moved closer to optimum. By using an automated tuning approach it would be possible to save manual work of finding hyperparameter, and it would guarantee that the model would be trained under optimal conditions, leading to improved predictive performance and generalization on the unseen set of data.

3.5 Evaluation

Following the hyperparameter optimization phase utilizing Optuna, the LightGBM model was retrained utilizing the best-performing hyperparameters on the training dataset. The finalized model was then evaluated on the test set to evaluate its predictive capability on unseen data. To ensure a comprehensive evaluation, various standard classification performance metrics were computed, including Accuracy, Precision, Recall, and F1-score.

In the classification problems, the models predictions are judged based on four essential states: true positives (TP), false positives (FP) , false negatives and true negatives (TN) (FN) [21].

Accuracy:- measures the ratio of whole correct predictions between everything predictions as show in Eq1 [22]:

$$Accuracy = \frac{TN+TP}{FP+TN+TP+FN} \tag{1}$$

Recall:- indicates the proportion of actual positive cases that were correctly identified as show in Eq2 [23]:

$$Recall = \frac{TP}{TP+FN} \tag{2}$$

Precision:- refers to the proportion of true positive predictions among all instances that were predicted as positive as show in Eq3 [24]:

$$Precision = \frac{TP}{FP+TP} \tag{3}$$

F1-score:- is the harmonic mean of recall and precision, and it balances the two as show in Eq14 [25]:

$$Recall = 2\frac{Recall*Precision}{Precision+Recall} \tag{4}$$

4. RESULTS AND DISCUSSION

To evaluate the proposed ML framework for the early detection of GDM a sequences of trials were achieved under two distinct train-test data split scenarios: 70% training with 30% testing and 80% training with 20% testing. The goal was to evaluate the model generalization performance with both moderate and larger training datasets. The LightGBM classifier was initially trained utilizing its default hyperparameter configuration, which served as the baseline. In the second stage, the model was retrained using a set of optimized hyperparameters obtained via the Optuna optimization framework, which systematically explored the hyperparameter space to maximize validation accuracy. Notably, the optimized model achieved a high accuracy of 91% with the (80-20) train-test split, confirming its effectiveness and reliability in predicting gestational diabetes. Across all tested scenarios, the results demonstrated a clear improvement in classification performance following the optimization process. The optimized LightGBM consistently outperformed the baseline configuration in terms of recall, precision, F1-score, and accuracy, as detailed in Table 2. This indicates that hyperparameter optimization is crucial for achieving robust and reliable performance, especially in health-related applications where early diagnosis is critical.

Table 2. Model Performance Before and After Optimization

Model	Precision	Recall	F1-score	Accuracy
Baseline LightGBM (30%-70%)	82%	83%	82%	83%
Baseline LightGBM (20%-80%)	81%	82%	81%	82%
After optimazed (30%-70%)	90%	90%	90%	90%
After optimazed (20%-80%)	91%	91%	91%	91%

The confusion matrices in Figures 3 and 4 show the baseline LightGBM model's performance under 70–30 and 80–20 splits, respectively, with moderate accuracy and visible class imbalance. In contrast, Figures 5 and 6 demonstrate a clear improvement utilizing the Optuna-tuned LightGBM model.

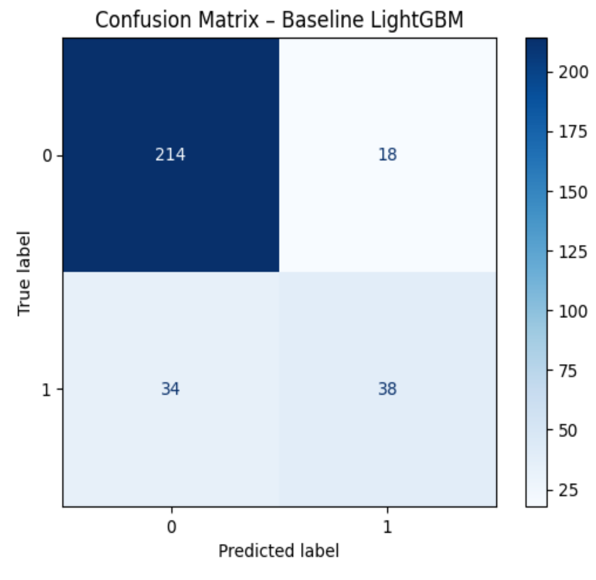

Figure 3. Baseline LightGBM (70-30)

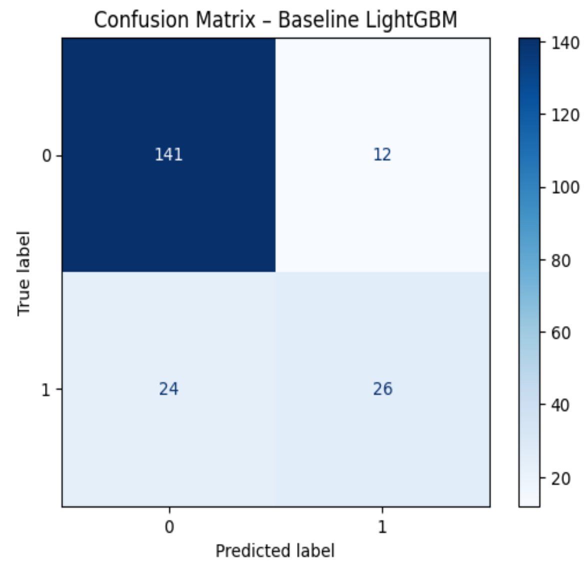

Figure 4. Base line LightGBM (80-20)

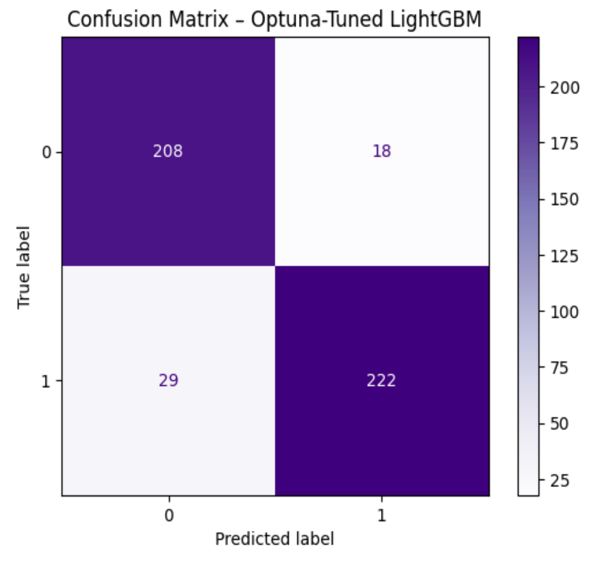

Figure 5. Optuna-Tuned LightGBM (70-30)

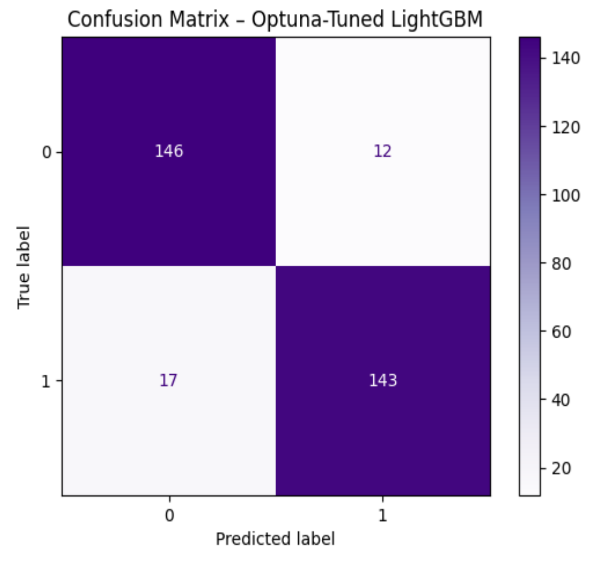

Figure 6. Optuna-Tuned LightGBM 80-20)

Table 3 shows a comparative analysis of the proposed model performance beside various state-of-the-art ML approaches reported in recent studies on GDM prediction. While traditional models like RF have shown moderate success achieving 80% accuracy in the study via Garg et al. [12] and 79% in the study via Bigdeli et al. [5] they fall short of higher benchmarks required for clinical reliability. Ensemble approaches such as Stacking, utilized via Zhao et al. [13] and Shetty et al. [15], reached 86.8%, marking a notable improvement over standalone models.

Other algorithms, including SVM, Extremely Randomized Trees, and RF generally yielded accuracies in the 74%–82% range across different datasets, reflecting some limitations in model generalization and data representation. In contrast, the proposed model, which integrates SMOTE for class balancing and Optuna for hyperparameter optimization within a LightGBM framework, achieved high accuracy of 91%, the highest among all compared studies.

Table 3. Comparation proposed model with related studies.

Reference	Year	Technique	Accuracy
Garg et al. [12]	2024	RF	80
		KNN	75%
		SVM	75%
		LR	77%
		DT	71%
		XGBoost	73%
Zhao et al. [13]	2024	RF	83.2%
		SVM	81.1%
		XGBoost	84.2%
		CatBoost	85.1%
		Stacking	86.8%
Bhuria et al. [14]	2024	LR	0.77%
Shetty et al. [15]	2024	LR	81%
		SVM	82%
		RF	80%
		GNB	77%
		XGBoost	80%
		VE	81%
		BDT	81%
		WAE	81%
		Stacking (XGBoost, NGBoost, AdaBoost)	81%
Kaya et al. [16]	2024	XGBoost	69.7%
		RF	65.15%
		LR	49.45%
		LGBM	55.05%
		AVG Blender	48.45%
		ET	74.25%
Victor et al. [17]	2024	LGBM	75%
		XGBoost	74%
		RF	76%
		CatBoost	75%
		AdaBoost	72%
Bigdeli et al. [5]	2025	RF	89%
		MLP	85%
		DT	82%
		NB	78%
		KNN	80%
		XGBoost	83%
		DT	70.16%
		RF	79.77%
		LGBM	76.66%
		XGBoost	81.05%
Proposed	2025	Optuna-Tuned LightGBM	91%

Despite the good results, this experiment underscores several important challenges. The approach successfully shows the dual benefits of combining robust hyperparameter optimization with effective data resampling to enhance classifier sensitivity and overall model performance. The model stays computationally efficient and interpretable, making it feasible for real-world clinical deployment. However, there are many limitations such as the dataset is relatively small and homogeneous, lacking the diversity needed for broader generalization. Moreover, the absence of longitudinal clinical data and a limited set of features reduce the models capacity to fully reflect patient variability. The confidence in synthetic data generation over SMOTE may introduce noise, and utilizing static patient snapshots fails to capture the dynamic changes that occur throughout gestation an important factor for timely and precise diagnosis.

5. CONCLUSION

This research proposed ML framework for prediction of GDM utilizing an optimized LightGBM classifier. Via combining SMOTE for class balancing and leveraging Optuna for hyperparameter optimization, the proposed model achieved a high accuracy of 91%, outperforming many existing techniques in recent literature. The model was evaluated under both 70–30 and 80–20 train-test splitting scenarios, consistently shown strong performance across key metrics like recall, precision, F1-score, and accuracy. The significant improvement over baseline models highlights the importance of optimization and data balancing in achieving clinical-grade prediction performance. Given its simplicity, efficiency, and high accuracy, the proposed approach holds strong potential for real-world clinical application, supporting healthcare professionals in the early detection and intervention of GDM. In Future work aims to enhance the model's robustness by utilizing a variety of larger-scale and longitudinal clinical notes, as well as integrating time-series data, and merging the model with other medical features. Advanced hybrid modeling approaches combining traditional ML (e.g., LightGBM, SVM) with DL (e.g., RNN, LSTM) will be explored. Further improvements include better feature selection, alternative class balancing methods like ADASYN, and the development of an interpretable clinical decision support system.

FUNDING INFORMATION

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

COFLICTS OF INTERESTS

The authors declare that there is no conflict of interest regarding the publication of this paper.

DATA AVAILABILITY STATEMENTS

The dataset used in this study is publicly available IEEE DataPort: "Gestational Diabetes" by Rasool Jader, https://ieee-dataport.org/documents/gestational-diabetes .

AUTHORS CONTRIBUATIONS

All authors contributed equally to this study design, implementation, data analysis, and manuscript writing and approved the final version.

EITHICAL APPROVAL

Ethical approval was not required for this work because it used a publicly available online dataset.

REFERENCES

[1] M. Niebrzydowska-Tatus et al., "Association of dpp-4 concentrations with the occurrence of gestational diabetes mellitus and excessive gestational weight gain," Int. J. Mol. Sci., vol. 25, no. 3, p. 1829, 2024.

[2] R. Modzelewski, M. M. Stefanowicz-Rutkowska, W. Matuszewski, and E. M. Bandurska-Stankiewicz, "Gestational diabetes mellitus—recent literature review," J. Clin. Med., vol. 11, no. 19, p. 5736, 2022.

[3] S. Suman, U. Pacharaney, and P. M. Gote, "A Type-1 Fuzzy Logic Framework for Early Risk Assessment of Gestational Diabetes Mellitus Using Clinical Risk Factors," in 2025 International Conference on Machine Learning and Autonomous Systems (ICMLAS), 2025, pp. 1180–1185.

[4] H. Y. Lu et al., "Digital health and machine learning technologies for blood glucose monitoring and management of gestational diabetes," IEEE Rev. Biomed. Eng., vol. 17, pp. 98–117, 2023.

[5] S. K. Bigdeli, M. Ghazisaedi, S. M. Ayyoubzadeh, S. Hantoushzadeh, and M. Ahmadi, "Predicting Gestational Diabetes Mellitus in the first trimester using machine learning algorithms: a cross-sectional study at a hospital fertility health center in Iran," BMC Med. Inform. Decis. Mak., vol. 25, no. 1, p. 3, 2025.

[6] J. Xing et al., "Enhancing gestational diabetes mellitus risk assessment and treatment through GDMPredictor: a machine learning approach," J. Endocrinol. Invest., vol. 47, no. 9, pp. 2351–2360, 2024.

[7] Y. Belsti et al., "Comparison of machine learning and conventional logistic regression-based prediction models for gestational diabetes in an ethnically diverse population; the Monash GDM Machine learning model," Int. J. Med. Inform., vol. 179, p. 105228, 2023.

[8] R. Jader and S. Aminifar, "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., vol. 2022, no. 1, p. 9749579, 2022.

[9] G. Cubillos et al., "Development of machine learning models to predict gestational diabetes risk in the first half of pregnancy," BMC Pregnancy Childbirth, vol. 23, no. 1, p. 469, 2023.

[10] N. Saleh, K. M. Yacoub, and A. M. Salaheldin, "Machine Learning_Based Paradigm for Diagnosis of Gestational Diabetes," in 2023 3rd International Conference on Electronic Engineering (ICEEM), 2023, pp. 1–4.

[11] S. R. Alapati et al., "Enhancing Gestational Diabetes Prediction Through Deep Learning and Traditional ML Techniques," in 2024 First International Conference for Women in Computing (InCoWoCo), 2024, pp. 1–7.

[12] S. Garg, S. Kumar, S. K. Singh, S. Kumar, V. Arya, and B. B. Gupta, "Gestational Diabetes Prediction using Machine Learning for Consumer Electronics Healthcare," in 2024 IEEE International Symposium on Consumer Technology (ISCT), 2024, pp. 243–249.

[13] K. Zhao, Z. Wang, and S. Guo, "Application of Ensemble Learning Model in Predicting Gestational diabetes," in 2024 13th International Conference of Information and Communication Technology (ICTech), 2024, pp. 57–63.

[14] R. Bhuria, K. S. Gill, S. Malhotra, and M. Singh, "Enhanced Diagnosis of Gestational Diabetes through Machine Learning and Logistic Regression," in 2024 4th Asian Conference on Innovation in Technology (ASIANCON), 2024, pp. 1–4.

[15] N. P. Shetty, J. Shetty, V. Hegde, S. D. Dharne, and M. Kv, "A machine learning-based clinical decision support system for effective stratification of gestational diabetes mellitus and management through Ayurveda," J. Ayurveda Integr. Med., vol. 15, no.

6, p. 101051, 2024.
[16] Y. Kaya, Z. Bütün, Ö. Çelik, E. A. Salik, T. Tahta, and A. A. Yavuz, "The early prediction of gestational diabetes mellitus by machine learning models," BMC Pregnancy Childbirth, vol. 24, no. 1, p. 574, 2024.
[17] A. Victor et al., "Predictive modeling of gestational weight gain: a machine learning multiclass classification study," BMC Pregnancy Childbirth, vol. 24, no. 1, p. 733, 2024.
[18] Rasool Jader, "Gestational Diabetes," IEEE Dataport. [Online]. Available: doi:10.21227/85ty-jf65 .
[19] N. Zamri et al., "A comparison of unsupervised and supervised machine learning algorithms to predict water pollutions," Procedia Comput. Sci., vol. 204, pp. 172–179, 2022.
[20] H. Tao et al., "Development of integrative data intelligence models for thermo-economic performances prediction of hybrid organic rankine plants," Energy, p. 130503, 2024.
[21] L. Lavazza and S. Morasca, "Common problems with the usage of f-measure and accuracy metrics in medical research," IEEE Access, vol. 11, pp. 51515–51526, 2023.
[22] Y. Jiao and P. Du, "Performance measures in evaluating machine learning based bioinformatics predictors for classifications," Quant. Biol., vol. 4, pp. 320–330, 2016.
[23] N. Mehdiyev, D. Enke, P. Fettke, and P. Loos, "Evaluating forecasting methods by considering different accuracy measures," Procedia Comput. Sci., vol. 95, pp. 264–271, 2016.
[24] P. Galdi and R. Tagliaferri, "Data mining: accuracy and error measures for classification and prediction," Encycl. Bioinforma. Comput. Biol., vol. 1, pp. 431–436, 2018.
[25] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," BMC Genomics, vol. 21, pp. 1–13, 2020.