

A Hybrid Diffusion-Augmented Ensemble and RAG-LLM Framework for Early Prediction of Type 2 Diabetes Mellitus

¹Sangeeta Lalwani, ²Dr. Shashi Kumar Singh

¹Research Scholar, Future University, Bareilly, India, sangeeta.lalwani29@gmail.com

²Dean (Engineering), Future University, Bareilly, India, kumar.shashi651@gmail.com

HOW TO CITE:

Sangeeta Lalwani, Shashi Kumar Singh (2026). A Hybrid Diffusion-Augmented Ensemble and RAG-LLM Framework for Early Prediction of Type 2 Diabetes Mellitus. International Journal of Special Education, 41(19s), 408-418

COPYRIGHT STATEMENT:

Copyright: © 2026 Authors.
Open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

ABSTRACT:

One of the most prevalent metabolic disorders that poses a challenge to healthcare systems due to its rising frequency and accelerating prevalence is Type 2 Diabetes Mellitus (T2DM) and consequently prediction and intelligent management is necessary to forbid the emergence and progression of disease at an early stage. The disordered physiology of T2DM is characterized by the interrelated coupling of genetic predisposition, clinical characteristics, behavioral patterns, and environmental exposures and thus requires the implementation of advanced predictive modeling frameworks that support early disease identification, patient risk assessment, and individualized therapeutic interventions. Although accuracy and reliability in diabetic prediction and management have been enhanced through legacy technologies such as Artificial Intelligence (AI), exceptionally Machine Learning (ML) and Deep Learning (DL), they are still prone to limitations that compromise intelligent clinical decision support framework. To mitigate the identified challenges, the proposed conceptual framework includes AI-driven architecture that synergizes diffusion-based Generative Artificial Intelligence, ensemble techniques, Explainable Artificial Intelligence (XAI), and Retrieval-Augmented Generation (RAG)-enabled Large Language Models into a unified decision support system for early prediction and personalized management of T2DM.

Keywords: Type2 Diabetes Mellitus (T2DM), ensemble learning, risk stratification, Retrieval Augmented Generation, Large Language Model.

1. Introduction

Type 2 Diabetes Mellitus (T2DM) is a global public health and economic challenge. Early prediction of T2DM is crucial for facilitating timely clinical intervention, mitigating disease progression, optimizing clinical decision support, and minimizing disease onset and its associated complications. Recent advances in Artificial Intelligence (AI) have facilitated robust predictive

frameworks capable of identifying complex, nonlinear relationships within heterogeneous clinical datasets. Integrating these predictive models with explainable artificial intelligence (XAI) techniques enhances model transparency and clinical interpretability, while personalized recommendation systems and evidence-based clinical guidelines support context-aware decision-making and individualized disease management.

Collectively, these AI-driven approaches have the potential to improve early risk stratification, promote precision medicine, and enhance overall clinical outcomes in T2DM management. [3]

Existing predictive frameworks predominantly exhibit black-box characteristics, limiting model interpretability, clinician acceptance, and their effective deployment in evidence-based clinical workflows. Furthermore, traditional Large Language Models (LLMs) generate responses exclusively from pretrained parametric knowledge, which may result in outdated information, hallucinated content, and inadequate grounding in authoritative medical evidence. [3] Initially, a diffusion-based generative model TabSyn is adopted to fabricate statistically representative tabular clinical data, thereby strengthening the quality of data. TabSyn performs on tabular data and in latent space overcoming the limitations of Synthetic Minority Over-sampling Technique (SMOTE) model that interpolates among the unrealistic data, generates unrealistic responses and thus cannot perform well on complex datasets. The study is followed by stacking ensemble learning framework that combines disparate base learners with a meta-learner to harness the complementary strengths of multiple predictive models resulting in reliable and trustworthy patient risk stratification. [19] To improve clinical explainability of model, Explainable AI techniques, including SHAP (Shapley Additive explanations) and LIME (Local

Interpretable Model-agnostic Explanations), are integrated in order to estimate feature importance to explain individual prediction outcomes. [20]

In addition to disease prediction the proposed framework also incorporates, Retrieval-Augmented Generation (RAG) with domain-specific Large Language Models to deliver empirically supported patient-centric healthcare. Before generating responses, the RAG module extracts relevant information from published biomedical research, authoritative medical databases and validated medical knowledge sources rather than relying on pretrained knowledge unlike LLM. This mechanism of extraction considerably reduces erroneous content generation, ensures factual coherence, integrates the latest medical guidelines, and provides profile driven recommendation to patient. Overall, the proposed framework elevates the conventional focus of disease prediction by integrating synthetic sample generation for data augmentation, stacking ensemble learning approach, comprehensible machine learning, context-grounded generative AI, and intelligent clinical decision support within a unified healthcare intelligence framework. The hybrid framework not only enhances the accuracy in disease prediction and management but also provide personalized recommendations and guidelines to patients thereby guiding them to make necessary modifications in their lifestyle habits regulating the progression of disease at an early stage. [5]

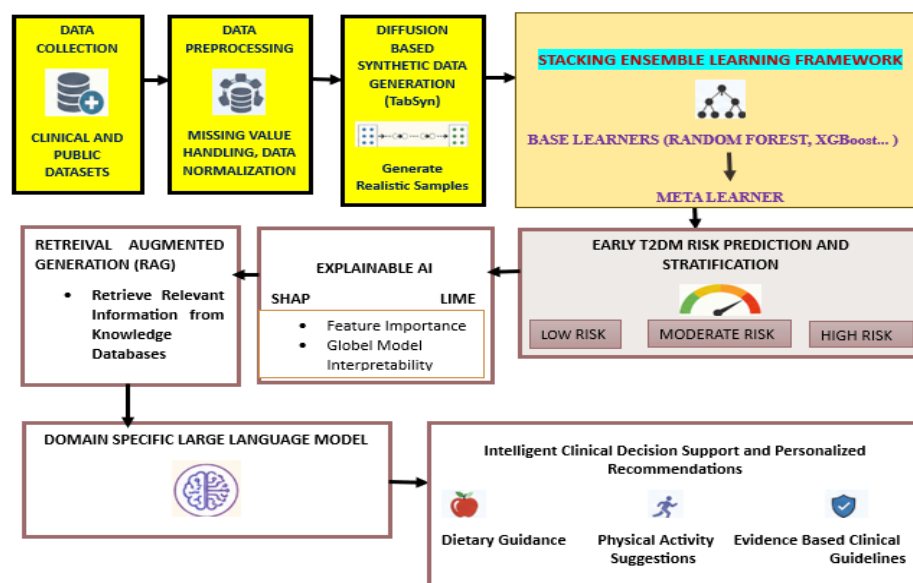


Fig. 1. Workflow for T2DM Prediction and Recommendation

I. LITERATURE REVIEW

Based on the existing literature, the research gap identified that there is no comprehensive framework that simultaneously integrates latent diffusion-based synthetic data generation, stacking ensemble learning, explainable AI, Retrieval-Augmented Generation, and Large Language

Models for early prediction, diabetic risk stratification, explainable decision-making, and personalized management of Type 2 Diabetes Mellitus. Therefore, the study proposes a unified hybrid framework that addresses these limitations while emphasizing prediction accuracy, explainability, trustworthy recommendations, and ethical AI-driven healthcare. [3]

Author(s) & Year	Key Contribution	Research Gap
Iparraguirre-Villanueva et al. (2023)	Involve various machine learning approaches such as Logistic Regression, Decision Tree, Random Forest, Support Vector Machines (SVM), K-Nearest Neighbour for prediction and management of diabetes at an earlier stage	Dependent on unbalanced datasets due to which they cannot provide with reliable predictions when employed on complex datasets.
Patro et al. (2023)	Deep learning architectures, such as ANN, CNNs, RNNs, have been employed in previous existing studies	High processing overhead is required and they function as black boxes as they provide the patient or user with early predictions but does not explain why such prediction was being made.
Pezoulas et al. (2024)	Adopted data enrichment methods such as SMOTE, Generative Adversarial Networks (GANs), Denoising	Lack in generation of realistic data and inefficiency in handling tabular data

	Diffusion Probabilistic Models (DDPM), and TabDDPM.	
Sampath et al. (2024)	Integration of Ensemble learning techniques such as Bagging, Boosting, Random Forest, XGBoost, AdaBoost, LightGBM, and CatBoost, have enhanced the reliability and accuracy in early prediction of diabetes	Lack proper risk stratification among the patients.
Oliullah et al. (2024)	Ensemble learning approach is found to be most suitable among the various AI approaches as it combines various base learners and predict the accuracy through majority voting and reduces bias-variance trade off	Explainability and personalized recommendations to patients is still lacking.
Sadeghi et al. (2024)	SHAP and LIME are some of the Explainable AI approaches that played a significant and vital role in decision making as they explain which diabetic factors contributed the most in causing T2DM	Still personalized recommendations are limited thereby functioning as black boxes leading to underdeveloped decision support system.

Al Nazi and Peng (2024)	Large Language Models (LLMs) have shown significant potential in providing patient or user with personalized recommendations and guidelines in order to make necessary modifications required in their lifestyles	Based on pretrained knowledge and are prone to provide outdated or inaccurate medical information
Gao et al. (2024)	Surveyed Retrieval-Augmented Generation (RAG) frameworks and demonstrated their ability to improve LLM based response by first extracting the information through semantic search from knowledge databases.	Decision support systems in development today lack this integration and therefore proper recommendations are not provided to users or patients.
Campanella et al. (2024)	Reviewed AI-based approaches for suggesting guidance and personalized recommendations to patients suffering from T2DM thereby developing a clinical decision support system	Lack of integration of ensemble learning approaches with explainable and LLM models thereby inability to provide patients with not only accurate prediction but also lack of personalized recommendations and effective management of T2DM.

Explainable AI Bibliometric Review (2026)	Identified growing adoption of XAI in diabetes risk prediction.	Limited work on personalized management.
Recent Generative AI Studies (2025–2026)	Demonstrated AI-generated recommendations and decision support in healthcare.	Very limited application to T2DM risk stratification and management.

II. RESEARCH METHODOLOGY

Through synergising the TabSyn diffusion based generative model, the proposed research methodology addresses the limitations of existing Type 2 Diabetes Mellitus (T2DM) prediction and management framework. In contrast to conventional oversampling techniques, TabSyn operates within a learned latent feature representation, promoting accurate representation of complex while preserving the multivariate dependencies. Furthermore, TabSyn incurs lower computational cost than traditional diffusion-based tabular generators while retaining quality of data, that can be effectively utilized for subsequent analysis and processing. [19]

TabSyn is a diffusion model mainly used to equalize the representation of diabetic and non-diabetic classes so that the machine learning model can process data effectively. The data is subsequently projected into a latent space representation to reduce computational complexity while preserving meaningful feature relationships. The data is progressively perturbed by Gaussian noise through a forward diffusion process. The model subsequently performs the reverse diffusion process to reconstruct high-fidelity synthetic patient records from the noisy latent representation. The model is designed to accommodate heterogeneous tabular data, thereby overcoming the limitations of conventional approaches such as numerical values (age, glucose level, BMI) as well as categorical values (gender, family history). The records are clinically meaningful as it accelerates model training by operating in the latent space, reduces memory consumption, and generates high-fidelity synthetic

records. The study adopts a stacking ensemble learning approach that leverages the complementary strengths of multiple base learners, resulting in enhanced predictive accuracy and generalization performance in prediction of diabetes and risk stratification. If $P_1, P_2, P_3, \dots, P_n$, are the predictions provided by the base learners and provided to meta-learner then the final prediction computed through meta-learning function $f(\cdot)$ is given as:

$$P_{\text{final}} = f(P_1, P_2, P_3, \dots, P_n)$$

In case meta-learner corresponds to logistic regression then the final prediction is expressed as:

$$P_{\text{final}} = \beta_0 + \beta_1 P_{\text{RF}} + \beta_2 P_{\text{XGB}} + \beta_3 P_{\text{LGBM}} + \beta_4 P_{\text{CatBoost}}$$

where β_i are learned coefficients that determine the contribution of each model.

The effectiveness of the stacking ensemble is explained by the bias–variance trade-off, which enables improved predictive performance through an optimal balance between underfitting and overfitting. Explainable AI techniques, namely SHAP and LIME, are integrated into the proposed framework to provide transparent and interpretable insights into the contribution of each predictor to the onset and progression of Type 2 Diabetes Mellitus. [22]

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|! (n - |S| - 1)!}{n!} [f(S \cup \{i\}) - f(S)]$$

- S is a subset of all features

- Nis the complete feature set.

This measures the average marginal contribution of feature i .

The proposed model following explainable AI techniques than integrates LLM (Large Language Model) along with RAG (Retrieval-Augmented Generation) to generate clinically relevant recommendations by leveraging the latest peer-reviewed medical literature. The mathematical formulation of the RAG framework integrated with LLM is expressed as follows: [5]

$$R = \text{Retrieve}(Q)$$

Where Q is user query and R is the response generated by RAG through investigating the

diabetic patient records through the current sources, medical research, online databases etc [3].

The final response becomes:

$$\text{Answer} = \text{LLM}(Q, R)$$

The response generated by RAG is given to LLM that processes the user query Q along with response R . The proposed framework generates personalized clinical recommendations by leveraging current medical evidence and established diabetes management guidelines, empowering individuals to implement appropriate lifestyle modifications, dietary adjustments, and physical activity regimens for effective prevention and management of Type 2 Diabetes Mellitus. [9]

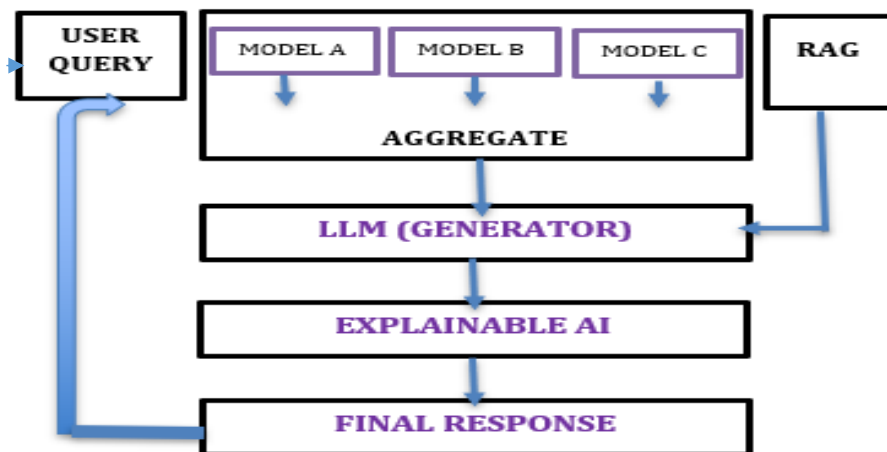


Fig .2. Conceptual Framework of the Model

III. Result Analysis

T2DM involves the use of DDPM, SMOTE, and TabDDPM for synthetic and high-quality data generation but still they face challenges when dealing with categorical and numerical datasets. Existing studies employed models that have several limitations. SMOTE (Synthetic Minority Over-sampling Technique) generate samples by interpolating between neighbouring instances that are of minority class and may be unrealistic. TabDDPM (Tabular Denoising Diffusion Probabilistic Model) is effective but may fail to capture the complex relationships between variables and does not focus on risk stratification studies. A powerful AI algorithm, Ensemble learning combines the predictions made from the

various base learners providing with more stable, reliable and accurate prediction along with risk stratification of diabetic patients. [17]

As it has been observed that medical datasets consist of noise, complex and missing values therefore the integration of ensemble techniques in prediction and management of T2DM plays a vital role in the study. Bagging, Boosting, Decision trees, Random Forest, XGBoost, AdaBoost, LightGBM, and CatBoost are some of the ensemble learning techniques widely used. Though Bagging methods improves the stability of the model by reducing variance and boosting methods improve the accuracy of prediction by reducing bias, the more powerful ensemble technique is Stacking as it minimizes the bias-variance trade-off through

meta-learning. Stacking Ensemble Learning optimally integrate the predictions of multiple base learners such as Random Forest, XGBoost, AdaBoost, Support Vector Machines(SVM) and Neural Networks and is considered as the most effective ensemble learning technique. Meta-learner is a high-level model that processes the outputs received by these base learners and provides with a final prediction of diabetic risk to the patient. The risk stratification using stacking ensemble learning approach is followed by objective to provide meaningful insights about the factors that influence the onset of the disease and then to provide recommendations through generative based models so that the precautions could be taken by the patients earlier before the onset of the disease. This is because a patient at high risk wants to know why he is being classified as such. [22]

Through the integration of SHAP and LIME into the developing model, a patient can know which factors of diabetes such as BMI, glucose level, insulin level, family history, blood pressure, contributed the most in categorizing the patient as high risk or low risk. Thus, the integration of explainable AI techniques into the model can clearly identify and explain the factors contributing the most on the disease onset and clearly explains the risk stratification being made by the model among the patients as high risk and low risk. This makes the prediction provided by the stacking ensemble learning techniques as explainable, personalized and trustworthy. After the risk stratification, prediction and explanation of factors about the T2DM disease been made using the stacking technique based on ensemble learning, there is a need for providing personalized recommendations to patients so that they can recover at an early stage before the onset of disease. For this, the model integrates LLM with RAG. The limitation with LLM is that it does not depend on

latest research findings and medical guidelines and therefore it can provide with recommendations based only on the previous studies and findings. RAG overcomes this limitation by integrating LLM with recent diabetic guidelines, recommendations from WHO, health records available online, patient's records and databases. RAG first retrieves appropriate information from these sources and then provides it to LLM so that it generates semantically informed response to patients. The existing studies include LLM that responds based only on the knowledge being trained to the model. However, integrating RAG the research focuses on developing the model such that the information is retrieved from the relevant databases and sources such as diabetic guidelines, recent studies and patient records which then LLM utilizes to provide the diabetic patients with personalized recommendations and guidance on modification of lifestyle, updating their dietary plans, physical activities required on regular basis etc. This helps in reducing the onsets of disease and its progression. RAG when integrated with LLM is capable of generating meaningful insights about early prediction, progression and regulation of Type2 Diabetic Mellitus (T2DM) as it has the capability to transform the predictions being made numerically into understandable form such that the patient could understand about the disease onset at an earlier stage and can take necessary precautions. Although the predictive performance of the model is evaluated through accuracy, additional evidence of the model's capability is provided through Precision, Recall and Specificity supporting the model to accurately distinguish between diabetic and non-diabetic individuals. Thus, these metrics collectively supported a comprehensive performance assessment for the early prediction and risk stratification of Type 2 Diabetes Mellitus. [20]

TABLE I. PREVIOUS STUDY VS PROPOSED FRAMEWORK

PREVIOUS STUDY LIMITATION	PROPOSED FRAMEWORK
SMOTE unrealistic samples;	DDPM, SMOTE, TabDDPM: realistic synthetic data; numerical & categorical

TabDDPM limited complex relationships/risk stratification.	handling; supports risk stratification.
Single models, bagging and boosting have bias/variance limitations.	Stacking ensemble with RF, XGBoost, AdaBoost, SVM, and meta-learner minimizes bias-variance.
Black-box predictions with limited interpretability.	SHAP and LIME explain feature contributions (BMI, glucose, insulin, BP, family history).
LLMs may use outdated knowledge.	RAG retrieves latest guidelines, WHO, studies and patient records.
Limited actionable outcomes.	Personalized recommendations, early prediction and risk stratification.

IV. LIMITATIONS

Although the integration of TabSyn Generative Diffusion Model, stacking ensemble learning, SHAP, LIME, RAG and LLMs in developing model have substantially enhance the early prediction and management of Type2 Diabetes Mellitus (T2DM) thereby providing tailored recommendations to patients, the study still has some limitations [3]. The results of prediction may be imprecise when the dataset exhibits missing values, is incomplete, noisy, or does not encompass diverse patient populations. In addition, the model may demonstrate limited predictive effectiveness when tested on datasets from different countries, hospitals and regions. [3]

Explainable AI techniques SHAP and LIME integrated into the framework may generate low-fidelity explanations when diverse feature engineering and learning techniques are incorporated in computation of feature importance. RAG-LLM integration may synthesize responses lacking evidential support in case the semantic search from knowledge databases is of inferior, inconsistent and lacks accuracy. The model

does not include domain-specific information, such as medical images, wearable device data, lifestyle habits of patients thereby compromising the accuracy of predictions and recommendations to users or patients. Finally, the framework has been evaluated only in an experimental environment and has not yet been tested in real hospitals. Before it can be used in routine healthcare, further studies are needed to evaluate its performance, security, privacy, scalability, and compatibility with existing hospital systems.

V. FUTURE SCOPE

The model can be further evaluated in future using diverse datasets with high quality from different hospitals, regions and countries in order to improve its prediction accuracy, reliability and to cover different populations. The future research should include both structure and unstructured data such as medical images, wearable devices, lifestyle habits and dietary plans so that better recommendations could be provided and precautions could be taken by patient or user at an early stage before the progression of disease. Through regular updating of RAG Knowledge base, the semantic search

could be made more effective as personalized recommendations could be provided to patients based on latest clinical guidelines and diabetic reports. This will also reduce the processing time and complexity of the advanced techniques being integrated into the model. In order to provide a healthcare patient with reliable, consistent, clear and more personalized recommendations, the future study should focus on more advanced explainable AI techniques. The ethical standards, privacy and security should also be integrated so that the datasets of patients could be safe and secure.

VI. CONCLUSION

This research focused on developing an effective framework for the early prediction of Type 2 Diabetes Mellitus (T2DM) by combining several artificial intelligence techniques into a single system. The study addresses limited data availability and class imbalance through integrating TabSyn diffusion model, which produced synthetic instances that preserve complex inter-feature dependencies among the diabetes related clinical parameters thereby improving the data quality and fidelity. The model leverages stacking ensemble

learning approach thereby improving the prediction accuracy by aggregating the strengths of heterogeneous machine learning algorithms. The proposed framework achieved elevated accuracy in risk prediction and demonstrated effective risk stratification. In addition, the model also quantified the contribution of diabetes-related clinical parameters to disease prediction by augmenting explainable AI techniques SHAP and LIME. These explanations enhance model transparency and trustworthiness by mitigating the black-box nature of the predictive model. The research also contributes in providing patient-centric and personalized recommendations through the aggregation of RAG with LLM. The model generates comprehensible clinical guidance through semantic retrieval from recent scientific literature and biomedical databases facilitating evidence-based lifestyle modifications supporting timely adoption of individualized lifestyle, dietary, and physical activity interventions. Thus, the practical applicability of the proposed framework is enhanced by unifying disease prediction, explainable decision-making, and personalized clinical decision support within a single intelligent system.

VII. REFERENCES

- Iparraguirre-Villanueva, O., Espinola-Linares, K., Flores-Castañeda, R. O., & Cabanillas-Carbonell, M. (2023). Application of machine learning models for early detection and accurate classification of type 2 diabetes. *Diagnostics*, 13(14), Article 2383. <https://doi.org/10.3390/diagnostics13142383>
- Sampath, P., Elangovan, G., Ravichandran, K., et al. (2024). Robust diabetic prediction using ensemble machine learning models with synthetic minority over-sampling technique. *Scientific Reports*, 14, Article 28984. <https://doi.org/10.1038/s41598-024-78519-8>
- Mackenzie, S. C., Sainsbury, C. A. R., & Wake, D. J. (2024). Diabetes and artificial intelligence beyond the closed loop: A review of the landscape, promise and challenges. *Diabetologia*, 67(2), 223–235. <https://doi.org/10.1007/s00125-023-06038-8>
- Jacobs, P. G., Herrero, P., Facchinetti, A., et al. (2024). Artificial intelligence and machine learning for improving glycemic control in diabetes: Best practices, pitfalls and opportunities. *IEEE Reviews in Biomedical Engineering*, 17, 19–41. <https://doi.org/10.1109/RBME.2023.3331297>
- Campanella, S., Paragliola, G., Cherubini, V., Pierleoni, P., & Palma, L. (2024). Towards personalized AI-based diabetes therapy: A review. *IEEE Journal of Biomedical and Health Informatics*, 28(11), 6944–6957. <https://doi.org/10.1109/JBHI.2024.3443137>
- Rahman, M. S., Rahman, M. A., & Hossain, M. A. (2024). Efficient machine learning approach for diabetes mellitus disease prediction. In *Proceedings of the 6th International Conference on Pattern Analysis and Intelligent Systems* (pp. 1–6). <https://doi.org/10.1109/PAIS62114.2024.10541247>
- Vasudevan, A., & Gopalakrishnan, P. (2024). Diagnosing diabetes using machine learning-based predictive models. *Procedia Computer Science*, 233, 288–294. <https://doi.org/10.1016/j.procs.2024.03.218>
- Sharma, A. K., & Gupta, S. (2024). Navigating the digital landscape of diabetes care: Current state of the art and future directions. *Procedia Computer Science*, 237, 336–343. <https://doi.org/10.1016/j.procs.2024.05.113>
- American Diabetes Association Professional Practice Committee. (2025). Standards of care in diabetes—2025. *Diabetes Care*, 48(Supplement 1). <https://doi.org/10.2337/dc25-SINT>
- American Diabetes Association Professional Practice Committee. (2024). Standards of care in diabetes—2024. *Diabetes Care*, 47(Supplement 1). <https://doi.org/10.2337/dc24-SINT>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). Association for Computing Machinery. <https://doi.org/10.1145/2939672.2939785>

- Ke, G., Meng, Q., Finley, T., et al. (2017). LightGBM: A highly efficient gradient boosting decision tree. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 3146–3154).
- Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: Unbiased boosting with categorical features. In *Advances in Neural Information Processing Systems* (Vol. 31, pp. 6639–6649).
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), 1189–1232. <https://doi.org/10.1214/aos/1013203451>
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(2), 123–140. <https://doi.org/10.1007/BF00058655>
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321–357. <https://doi.org/10.1613/jair.953>
- Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems* (Vol. 33, pp. 6840–6851).
- Kotelnikov, A., Baranchuk, D., Rubachev, I., & Babenko, A. (2023). TabDDPM: Modelling tabular data with diffusion models. In *Proceedings of the 40th International Conference on Machine Learning* (Vol. 202, pp. 17564–17579). PMLR. <https://doi.org/10.48550/arXiv.2209.15421>
- Shaheen, S., et al. (2024). Hi-Le and HiTCL: Ensemble learning approaches for early diabetes detection using deep learning and explainable artificial intelligence. *IEEE Access*, 12, 66516–66538. <https://doi.org/10.1109/ACCESS.2024.3398198>
- Wang, X., Ren, J., Ren, H., et al. (2023). Diabetes mellitus early warning and factor analysis using ensemble Bayesian networks with SMOTE-ENN and Boruta. *Scientific Reports*, 13, Article 12718. <https://doi.org/10.1038/s41598-023-40036-5>
- Wolpert, D. H. (1992). Stacked generalization. *Neural Networks*, 5(2), 241–259. [https://doi.org/10.1016/S0893-6080\(05\)80023-1](https://doi.org/10.1016/S0893-6080(05)80023-1)
- Rokach, L. (2010). Ensemble-based classifiers. *Artificial Intelligence Review*, 33(1–2), 1–39. <https://doi.org/10.1007/s10462-009-9124-7>
- Dietterich, T. G. (2000). Ensemble methods in machine learning. In *Proceedings of the International Workshop on Multiple Classifier Systems* (pp. 1–15). Springer. https://doi.org/10.1007/3-540-45014-9_1