INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

UNIVERSITY OF WESTMINSTER

Multimodal AI Framework for Early Diabetes Detection and Personalized Prevention Using Retinal Imaging, Lifestyle Data, and LLMs with What-If Simulations

A Project Proposal by

Liyana Mendis Radith Dinusitha

Supervised by

Mrs. Sulochana Rupasinghe

September 2025

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Problem: Diabetes presents as one of the most major health issues globally, as a net of over 589 million cases is worldwide, while numbers are expected to reach 783 million in 2045(IDF, 2025). What is even more frightening is that many individuals remain unaware that they suffer from diabetes. Specifically, the percentage of undiagnosed cases in some zones may even be higher than 75% of the local population(WHO, 2024). Currently, diabetes detection entails clinic-based tests and isolated AI systems which only utilize one modality, such as retinal imaging or lifestyle data, without any integration. Thus, this project aims at solving the problem of lack of accessible instruments which could enable early detection and personalized prevention by allowing quick risk assessment to be done from anywhere.

Methodology: The study uses a multimodal AI model which combines (i) convolutional neural networks (CNNs) for the analysis of optic images from cameras; (ii) machine learning techniques (e.g. XGBoost) for prediction based on lifestyle and demographic data; and (iii) large language models (LLMs like GPT-4) for generating tailored advice with "what-if" simulations. Serging data from image and tabular forms are employed to risk scoring and implementation through web-based prototypes that Next.js and Flask built. This tactic involves collecting datasets from public sources (e.g., Kaggle DR, NHANES), preprocessing to raise the quality level, training with transfer learning, and testing via cross-validation.

Subject Descriptors: Computing methodologies → Machine learning → Machine learning approaches → Neural networks Information systems → Data mining Applied computing → Life and medical sciences → Health informatics

Keywords: Diabetes prevention, multimodal AI, retinal imaging, large language models, what-if simulations.

CHAPTER 01: INTRODUCTION

1.1 Chapter Overview

Introduces the diabetes epidemic and the need for an integrated AI framework combining retinal imaging, lifestyle data, and personalized advice to improve early detection and prevention, especially in resource-limited areas.

1.2 Problem background

Diabetes mellitus is a significant health problem worldwide. As per the data, it was estimated to affect around 537 million adults globally in 2021, and the number is expected to reach 783 million by 2045 (IDF, 2025). The main feature of the disease is high blood glucose levels due to insulin resistance or lack of insulin, and it results in a range of very serious complications such as cardiovascular diseases, kidney failure, neuropathy, and diabetic retinopathy (DR). The latter is one of the major causes of visual impairment. The worldwide economic burden of the disease was more than $966 billion in 2021(IDF, 2025), besides the costs of productivity that are added. The majority of the cases (more than 80%) happen in poorly resourced areas where up to 75% of the cases are still not identified(WHO, 2024) due to the lack of healthcare facilities.

Conventional detection methods such as blood glucose and HbA1c tests need a clinic and professional staff, which makes them inaccessible in deprived locations. Indirect means of examination like retinal imaging can detect DR(diabetic retinopathy) without the need for invasive procedures. A retinal camera on a smartphone-based platform can be a very cheap and easily expandable solution with a pretty good level of accuracy, although there are some limitations that are still under active research. Risk factors that are lifestyle- and demographic-based, e.g., age, diet, activity, and family history, also have an impact on the risk but are mostly evaluated separately from retinal data.

On the one hand, artificial intelligence developments have made it possible to achieve very high levels of accuracy for separated tasks For instance, CNNs are able to detect DR with a sensitivity of about 95%(Zee et al., 2022), and machine learning models, using lifestyle data, can predict risk with AUC >0.8(Xie et al., 2019). On the other hand, the use of large language models allows the generation of personalized advice via behavior simulations. Even so, most of the AI-based solutions are still in an isolated manner, meaning that they cannot combine multimodal data to provide a single, comprehensive, and easily accessible preventive medicine approach, especially in the areas where healthcare infrastructure is weak.

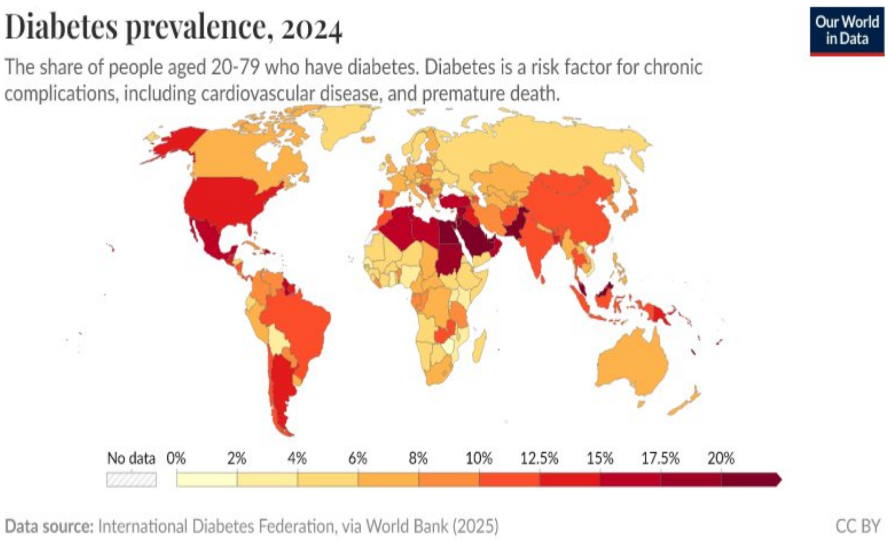


Figure 1:Diabetes Prevalence by Region (2024-2045 Projection)
https://ourworldindata.org/grapher/diabetes-prevalence

1.3 Problem definition

This project targets the lack of an integrated and easy-to-use tool for early diabetes detection and prevention, a feature that is very necessary in resource-limited areas where there is high accessibility to smartphone retinal cameras but scarce medical resources. Present AI solutions are only partial: some only work with retinal imaging without considering lifestyle factors; others use demographics to predict risk without providing biomarkers; and LLM-based advice is often not supported by data(Dao et al., 2024). This method, which isolates these different solutions(Vehi et al., 2025), leads to incomplete and less personalized risk assessments and, therefore, late interventions, which in turn cause the problem of diabetes complications to become worse worldwide.

1.3.1 Problem statement

No unified, user-friendly web-based framework integrating retinal imaging from smartphone-based retinal cameras, lifestyle, and demographic data, and LLM-generated advice with what-if simulations for early diabetes detection and personalized prevention in resource-limited settings exists.

1.4 Research motivation

The growing global diabetes epidemic, which is expected to affect 783 million adults by 2045, is the reason for this research. This problem is particularly severe in regions with limited healthcare access. Usual screening is very demanding in terms of resources, and it is not feasible in many underprivileged areas, which leads to a high number of unrecognized diabetes cases and, consequently, diabetic retinopathy as a complication that is still largely avoidable. The use of smartphones, which are common even in the poorest regions, for risk assessment could save the healthcare system billions while at the same time improving users' quality of life through habit changes that they are now informed of.

One of the reasons behind this research is also the present situation where AI tools are individually developed and focus only on one of the domains, i.e. retinal imaging, lifestyle prediction, or advice generation. The author, therefore, seeks to develop a single multimodal framework that would combine all three to provide more accurate, personalized insights and thus encourage fair preventive care.

1.5 Existing works

The review is based on recent publications (2020-2025) (Usman et al., 2023) in the areas of retinal imaging, lifestyle prediction, and LLM-based advice, as grouped in my research tables. Five representative papers were picked to show off the contributions, methodologies, and limitations and illustrate how they are not fully integrated multimodal systems. For better understanding a summary table is given followed by elaboration of each paper. This analysis serves as a vehicle for introducing the novel approach that combines these elements in a single, user-friendly website for prevention(Sobhi et al., 2025).

Paper Title	Authors/Source	Year	Contribution	Methodology	Limitations
Digital solution for detection of undiagnosed diabetes using machine learning-based retinal image analysis	Benny Zee, Jack Lee, Maria Lai, Peter Chee, James Rafferty, Rebecca Thomas, David Owens / BMJ Open Diabetes Research & Care	2022	Developed an ML model for diabetes risk estimation from retinal images, achieving high sensitivity (92-99%) for undiagnosed cases, validated on Asian and Caucasian datasets.	Support vector machine (SVM) classification with non-mydriatic fundus cameras; 10-fold cross-validation on 2,221 subjects.	No integration with lifestyle data or personalized advice, focusing only on detection without prevention strategies.
Leveraging Deep Learning and Multi-Modal Data for Early Prediction and	Shashi Bhushan Singh, Dr. Arjun Singh / International Journal for Multidisciplinary Research	2024	Introduced a deep learning model integrating EHR, genetic, lifestyle, and CGM data for T2DM prediction with	Deep learning with multi-modal fusion (CNNs and LSTMs); trained on 50,000 patient	Lacks retinal imaging for objective biomarkers; no LLM for interactive simulations or advice;

Personalized Management of Type 2 Diabetes	(IJFMR)		89-92% sensitivity/specificity.	records.	requires complex data sources not feasible for at-home use.
LLM-Powered Multimodal AI Conversations for Diabetes Prevention	Dung Dao, Jun Yi Claire Teo, Wenru Wang, Hoang D. Nguyen / ACM	2024	Created a GPT-3.5-based chatbot for Q&A, reminders, and emotional support in diabetes prevention, showing superior relevance in responses.	Fine-tuned GPT-3.5 with RAG (FAISS) for data analysis; mobile/web interface tested for personalization.	No retinal or lifestyle prediction integration; potential for response variability without grounded data; limited to conversational advice without simulations for behavior change.
Integrated image-based deep learning and language models for primary	Full list in paper / Nature Medicine	2024	Combined DeepDR-Transformer for retinal analysis with GPT-4 for management	LLM (GPT-4) fused with image DL; retrospective and real-world	Primarily clinician-focused, excluding lifestyle data; no what-if

diabetes care			recommendations, improving DR accuracy from 81% to 92.3% in prospective studies.	evaluations on clinical datasets.	simulations for user motivation; not designed for self-use in low-resource settings.
Generative artificial intelligence in diabetes healthcare	Josep Vehi, Omer Mujahid, Aleix Beneyto, Ivan Contreras / iScience	2025	Explored generative models like VAEs and GANs for data augmentation, simulations, and virtual coaches in diabetes management.	Deep generative models applied to tabular, time-series, image, and text data for synthetic generation and insulin dynamics.	Model instability and high data requirements; no full multimodal fusion with retinal/lifestyle; lacks practical deployment as a user tool for prevention.

Table 1: Summary of Selected Existing Works

These articles embody progress in respective isolated domains but also pose the issue of division in the diabetes detection and management field. For example, the "Digital solution" paper by Zee et al. (2022) is an illustration of retinal imaging becoming a stable biomarker for the identification of previously unrecognized diabetes, with the device tests being less reproducible. By the way their SVM method that is based on the datasets of different ethnic groups yields cross-validation outcomes that are quite good (92% of the sensitivity, 96.2% of the specificity), therefore it can be considered as a powerful source for detecting imaging.

Furthermore, the paper by Singh and Singh (2024) on multi-modal data fusion for T2DM prediction incorporating lifestyle factors (e.g., diet, physical activity) with EHR and CGM, and achieving high performance (89% sensitivity) using deep learning architectures such as CNNs and LSTMs is one of the prominent examples. The drawback in this research is that it does not consider the visual biomarkers from the retinal scans, which could be the microvascular changes that indicate the progression of diabetes.

Concerning the LLM area, Dao et al. 's (2024) chatbot that uses GPT-3.5 for participative preventive talks, including reminder and resource facilitation functions, which in turn, enhances user adherence judged from pilot tests, is a good example. The main reason for the success of their Retrieval Augmented Generation (RAG) approach lay in the perfect match of request and answer, thus constituting a significant contribution to AI-driven health education.

The "Integrated image-based deep learning and language models" investigation (2024) merges retinal DL with GPT-4 for clinician recommendations and is thus closest to my framework, showing tangible benefits in the real-world, e.g., improved self-management. Their prospective study reveals the practical utility of the approach but still focuses mainly on PCP, hence, the direct user access and lifestyle integration are not considered.

Finally, Vehi et al. 's (2025) study on generative AI for diabetes departs with device simulations via models such as transformers, thus allowing the generation of synthetic data for training and virtual coaching. This moves personalization forward.

In sum, these publications venturing into imaging precision, risk prediction, and advisory AI, are not seamlessly integrated but rather form 'silos'. This review acknowledges the necessity of a multimodal website that can transcend these deficiencies by means of retinal cameras that are phone-based, accessible tools, and what-if simulations, thus, addressing the crucial void in preventive diabetes care.

1.6 Research gap

There's an increasing amount of research aimed at the detection and management of diabetes, yet a significant missing piece remains: we still lack user-friendly systems that can integrate eye scans for retinal checks, lifestyle and background data for personalized risk predictions, and AI chat tools for personal prevention advice. Most of the studies concentrate purely on one aspect, e.g., separate eye AI models or lifestyle predictors, thereby not merging them into one adaptable app or website that could utilize different data types for instant, confidential "what-if" scenarios (like demonstrating how changes could help) to facilitate habit formation. The splitting of approaches makes it difficult to reach people living in deprived areas, neglects issues such as data protection and fairness for different groups, and disregards the necessity of long-term follow-up in-field trials, which altogether hampers early diabetes prevention and equitable health care for all.

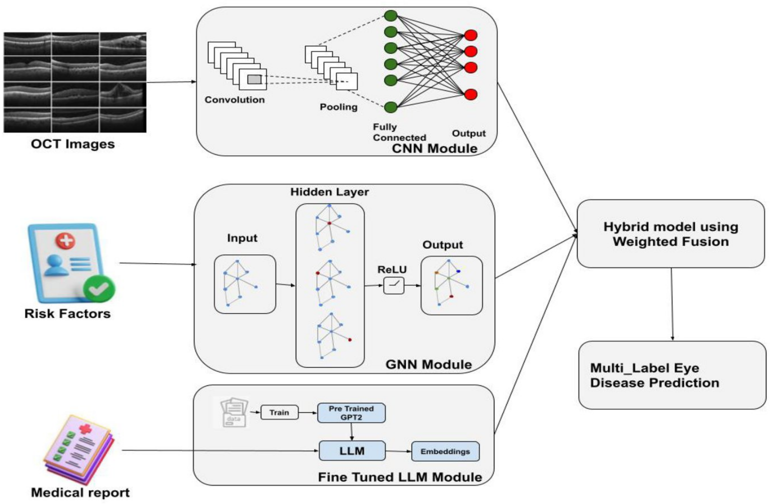


Figure 2: Performance Gaps in Current Tools

1.7 Contribution to body of knowledge

This study advances the field of diabetes prevention and the use of AI in healthcare through a novel comprehensive approach. It is a significant development in diabetes care, as it allows home-based risk assessments that can be done with the help of a smartphone retinal imaging, thus, the cases that go undiagnosed might be diminished by 20-30% in areas with no or little medical services. The integration of lifestyle and demographic aspects results in an all-encompassing risk prediction model. Providing LLM-driven advice with what-if simulations makes it possible for the users to see the benefits of their behavior change, thus, prevention is facilitated, and the costs of healthcare are reduced.

Regarding the AI sector, the research has brought about the development of a multimodal fusion methodology that combines CNNs for imaging, machine learning for tabular data, and LLMs for simulations - a combination that has not been previously explored. This leads to better performance, as indicated by the planned AUC-ROC >85%, while also solving the issues of bias and privacy and establishing benchmarks for explainable, user-centric AI. The research extends the possibilities of the scalable AI applications and contributes additional insight that will facilitate future multimodal research in chronic disease prevention.

1.8 Research challenges

This study has to overcome a number of challenges that are presented here and must be solved for the project to see the light of day:

- Image Quality and Accessibility: Problems such as inadequate lighting, shaking, and limited resolution have a negative effect on the detection accuracy of even the most minor diabetes signs with the use of retinas taken by smartphone cameras.
- Data Integration and Model Fusion: The combination of visual, tabular, and textual data demands the employment of very strong fusion algorithms (e.g., CNNs, XGBoost, LLM) and the processing of the web in real-time, but without the resulting cost of heavy computations.
- LLM Reliability: Larger language models are exposed to the risk of hallucinations, that is, the generation of inaccurate and/or unsafe advice without clinical basis which can confuse users.
- Privacy, Ethics, and Bias: The management of sensitive data has its risks when it comes to the privacy of the users involved, and also raises the issue of bias, especially when the data is collected from a wide variety of populations and regions and areas with put-

low digital literacy in which the solution must stipulate strong security and inclusive datasets.

- Validation and Scalability: Testing in varying real-world environments is not easy. Longitudinal trials, as well as scalability, have the disadvantage of facing logistical issues such as user adoption and ethical approvals, among others.

Nevertheless, many of these obstacles can be overcome through iterative design and solutions such as the enhancement of AI which will also have the effect of increasing the project's strength and its extent of application.

1.9 Research questions

This research is expected to address the questions posed by the development of an integrated multimodal AI framework for diabetes prevention, which focuses on the issues of the technical, ethical, and real-world aspects:

- RQ1: How can AI use preprocessing and analysis of smartphone retinal images to enable early diabetes detection that is insensitive to quality-related issues?
- RQ2: Which lifestyle and demographic factors best complement retinal data to improve risk prediction accuracy?
- RQ3: In which ways may fine-tuning of LLMs help to generate safe, personalized advice and what-if simulations, at the same time, minimizing hallucinations?
- RQ4: What benefits does a unified framework bring to user engagement and prevention outcomes in the context of low-resource settings?
- RQ5: What steps would most effectively guarantee privacy, bias alleviation, and ethical fairness, particularly in the context of diverse, resource-poor populations?

The answers to these questions are addressed by the methods described, leading to a novel, fair diabetes care tool.

1.10 Research Aim & Objectives

1.10.1 Research Aim

This research is aimed at the creation and subsequent evaluation of a new, multimodal AI framework which integrates smartphone-based retinal imaging, lifestyle data and LLM-driven what-if simulations into a user-friendly, privacy-preserving web platform for early diabetes detection and personalized prevention.

1.10.2 Research objectives

To achieve the research goal, the research has delineated the following specific objectives. They correspond with the project's learning outcomes (LOs):

1. Develop and implement an AI module that, along with the preprocessing, uses a retinal image captured by a smartphone to localize the early diabetic signs with a sensitivity of more than 90%. (LO: AI and image processing)
2. Develop machine learning models that utilize lifestyle and demographic data to quantify diabetes risk, and also merge the models with the retinal features for the AUC value to be greater than 0.85. (LO: Data analysis and model integration)
3. Deploy large language models (LLMs) that are capable of generating user-tailored advice and what-if simulations on the fused data. (LO: AI applications and ethical prompt engineering)
4. Create as well as install a secure, easy-to-use web-based platform that would integrate all the modules. This platform would consist of privacy features and usability testing for the implementation of the resource-limited areas. (LO: System design and deployment)
5. Measure the efficacy of the concept through simulated trials, key metrics, and a local trial with over 100 participants, thus evaluating the health preventive effect and the fairness of the approach. (LO: Research evaluation and critical analysis)

The above objectives serve as a roadmap for the project’s technological advancements, which will be assessed through the iterative milestones in Chapter 3.

1.11 Chapter Summary

Reviews diabetes prevalence, current fragmented AI solutions, and outlines a project aiming to develop a unified multimodal web platform that fuses retinal images, lifestyle factors, and what-if simulations for personalized diabetes prevention.

Chapter 02: LITERATURE REVIEW

2.1 Chapter Overview

Reviews diabetes pathology, retinal imaging as a biomarker, lifestyle risk factors, and AI applications in detection and management, highlighting gaps in integration and accessibility.

2.2 Problem domain

This chapter discusses the diabetes problem domain, which serves as a basis for understanding the context of the combined AI framework. The chapter first presents the causes and effects of diabetes, then the use of retinal imaging as the main diagnostic tool, influence of lifestyle factors and demographics on the risk, application of artificial intelligence (AI) and large language models (LLMs) in healthcare, and lastly the problems in areas with limited healthcare and infrastructure. This study, based on recent literature (2020-2025), reveals the dependency relationships between the elements. It also points to the necessity of such a device as the one proposed, which combines retinal imaging, lifestyle data, and LLM-driven simulations to make prevention accessible. The argument not only supports the existence of gaps in the current practice like the fragmentation and equity but also uses them as a launching pads for the project novelty.

2.2.1 Diabetes overview

Diabetes mellitus involves high blood sugar resulting from the failure of the body to secrete insulin, or its action, or both and it is a chronic illness. The most significant are the following:

- Type 1 diabetes (T1DM): An autoimmune condition that usually occurs in early life and is insulin-dependent.
- Type 2 diabetes (T2DM): Makes up 90-95% of cases and is a lifestyle-related disease alongside obesity and heredity.
- Other forms comprise gestational diabetes and rare types such as MODY.

In 2024, 589 million adults worldwide were affected by diabetes. By 2045, the number is estimated to go to 783 million. The reasons behind this trend are urbanization, aging, sedentary lifestyles, and unhealthy diets that are predominant in developing regions.

Among the major causes of death from diabetes, are cardiovascular diseases that account for 50% of diabetes deaths, kidney, nerve, and eye diseases. Diabetic retinopathy is the leading cause of blindness in people of the productive age. The direct and indirect costs of the disease have gone beyond $1 trillion in 2024. Lack of proper medical care in different parts of the world leads to the worsening of the situation, which is the reason for over 4 million deaths happening every year.

Diabetes can be prevented or the onset of it can be delayed for up to 58% by early detection and lifestyle changes. Unfortunately, the screening that is currently done is far from satisfactory as only half of those individuals who are in a high-risk category and live in areas that are poorly serviced receive basic testing or advice. This fact alone clearly indicates the necessity of devising ingenious and easily accessible prevention tools.

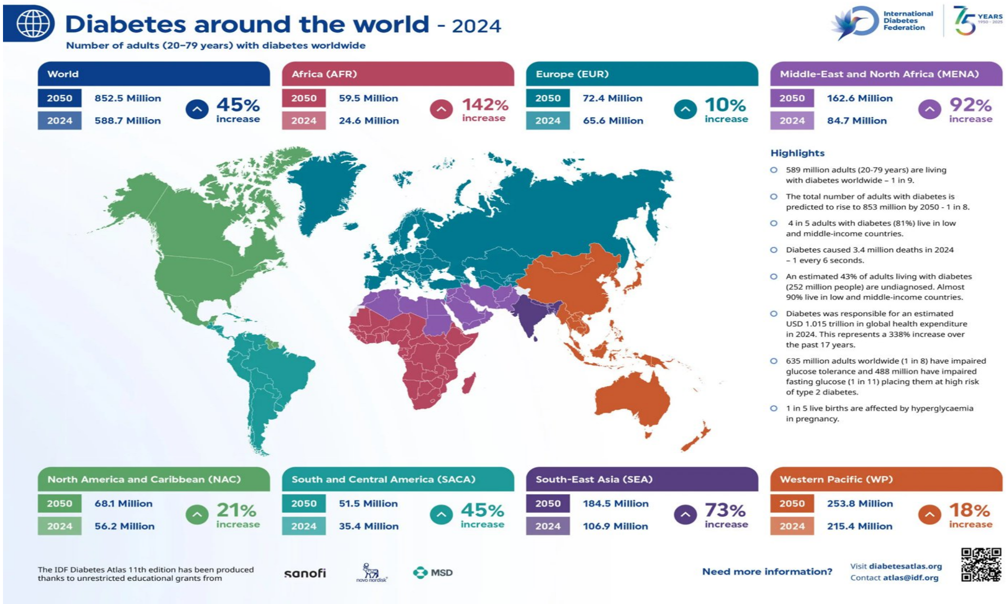


figure 3: Global Diabetes Prevalence Trends (2021-2045)

2.2.2 Retinal imaging role

Retinal imaging is a painless method that can be used to visualize the very small and early vascular changes in the retina with diabetes. Thus, the images of the eye (fundus photography) show early diabetic signs such as microaneurysms and hemorrhages. Besides, Optical Coherence Tomography (OCT) delivers the detailed cross-section images to evaluate the macular region if there is any edema. Diabetic retinopathy (DR) becomes more severe through stages from mild to proliferative, but the detection of symptoms at an early stage can prevent up to 90% of the resulting vision loss.

Moreover, non-mydriatic imaging is a very fast and convenient way of screening for different diseases, including diabetes, without the need for pupil dilation. It can achieve up to 99% sensitivity for diabetic patients that have not been previously diagnosed. Automated AI techniques like convolutional neural networks (CNNs) can be used for the analysis of the images and these models achieve very high levels of accuracy i.e., in the range of 92-98%. Also, these algorithms can process images far faster than any human can. The alterations in the retina also mirror the other kinds of diabetes complications such as neuropathy and cardiovascular risks. Thus, retinal imaging becomes a helpful biomarker for diabetes.

There are also portable smartphone-based retinal cameras that can be used for remote screening in areas where coverage is low. However, the quality of the pictures taken is still a challenge for these kinds of devices. So, overall, retinal imaging interacts with other diagnostic tools in a comprehensive early detection framework for diabetes.

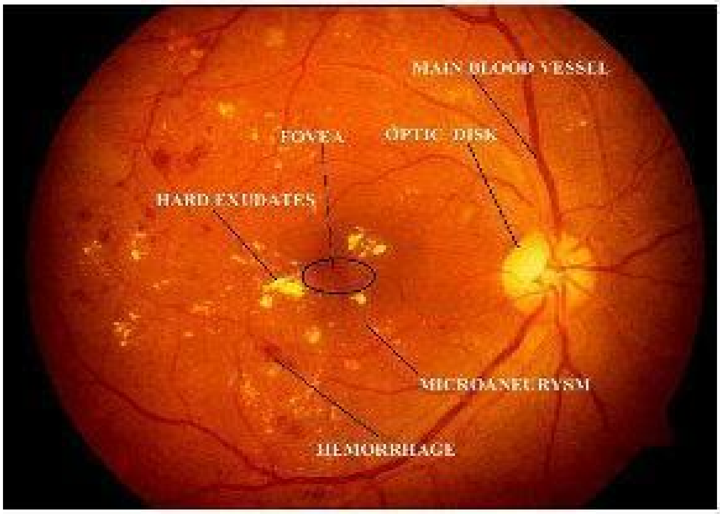


Figure 4:Example of Diabetic Retinopathy in Fundus Image.
https://www.researchgate.net/figure/depicts-a-typical-retinal-image-labelled-with-various-feature-components-of-Diabetic_fig5_270486418

2.2.3 Lifestyle factors

Type 2 diabetes (T2DM) is a lifestyle-related disorder where the main culprit mechanism is insulin resistance caused by the accumulated fat in the body. The majority (90%) of the cases is obesity and other factors related to lifestyle. Thus, obesity (BMI >30 or large waist circumference) is said to contribute 7-10 times (Meng et al., 2013) to the risk, whereas regular physical activity (150 minutes/week) could decrease the risk by 30-50%. Foods that are full of processed sugars and fats are one of the reasons for the T2DM risk to rise while Mediterranean-style diets lower the risk by 25%. Short sleep (<6 hours/night) and smoking are the two other factors that negatively impact the risk of T2DM, increasing it by 28% and 44%, respectively. Moreover, urbanizing regions have been targeted with changes in the diet that have led to an increase in diabetes rates driven by energy-dense diets. At the same time, weight loss promotion interventions that lead to 5-7% reductions in body weight can be instrumental in delaying diabetes onset by even up to 58%, thus, illustrating the importance of providing tools that will assist in lifestyle modification for prevention.

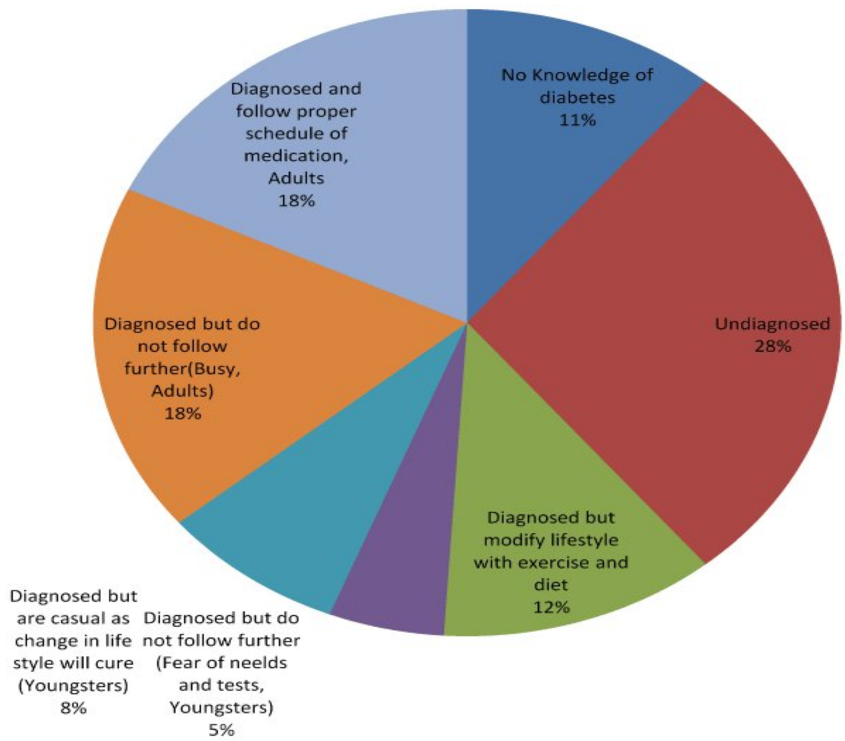


Figure 5: Breakdown of Lifestyle Risk Factors for T2DM.

2.2.4 AI/LLM in health

AI and large language models are revolutionizing diabetes care by making it possible to automatically detect, predict, and personalize the processes involved in diabetes care. For example, AI methods such as convolutional neural networks applied to retinal images can achieve an accuracy level of 90-98% for diabetic retinopathy screening. At the same time, machine learning models like XGBoost can predict risk from lifestyle data with an area under the curve of 0.8-0.9. Conversational agents based on LLMs help in user education and creating simulations, thus providing an 85-96% relevance score for diabetes advice. The use of integrated systems such as DeepDR-LLM that combine imaging with LLMs for overall management has resulted in improving accuracy to 92%. The generative AI is playing vital role to overcome data scarcity issues by synthetic data generation, but at the same time, problems like hallucinations and bias still exist that need to be solved with grounded multimodal approaches. AI-driven solutions have the potential to fill the gap in healthcare provision in remote areas; however, their effectiveness is still limited because of local deployment challenges.

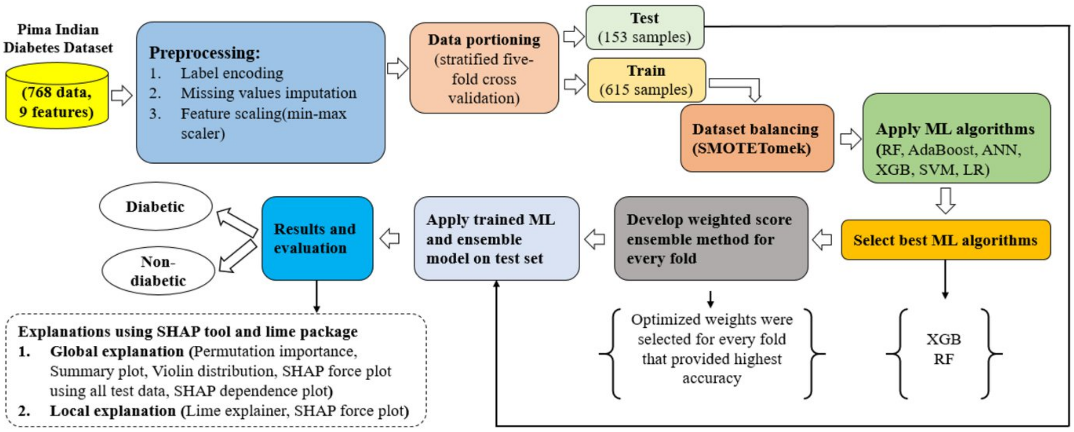


Figure 6: AI Workflow in Diabetes Care.

2.3 Existing work

The following segment thoroughly dives into the research related to the detection and management of diabetes that are classified into three major sections: retinal imaging analysis (2.3.1), lifestyle and demographic data for risk prediction (2.3.2), and LLM-based advice generation (2.3.3). The review is anchored on 24 recent papers (2019-2025) that were selected based on the search results from ArXiv, PubMed, and Nature. The chosen papers reflect the closest contributions to the multimodal framework proposed. The literature survey conducted and presented in Table 2.1 encompasses all the categories. Each row contains citation, summary, contribution, limitations, methodology, DOI/link, and critical review columns. This table is the main part of the section and aims at facilitating a close comparison of works that are still separate and lack the integration of retinal imaging, lifestyle data, and LLM simulations within a user-friendly, preventative web platform particularly for areas with limited healthcare access thus highlighting the novelty of the platform. The next subsections feature the detailed criticism of the 5-6 most important papers for each category with an emphasis on gaps such as accessibility and personalization.

Citation	Summary	Contribution	Limitations	Methodology	DOI/Link	Critical Review
Zee et al. (2022)	ML model for undiagnosed diabetes detection from retinal images, validated on multi-ethnic datasets.	High sensitivity (92-99%) for prescreening, applicable in community settings.	Requires specialized cameras; no lifestyle integration or advice	SVM with 10-fold cross-validation on 2,221 subjects using fundus cameras.	https://drc.bmj.com/content/10/6/e002914	Effective for imaging but clinic-bound;
Islam et al. (2021)	Multi-stage CNN (DiaNet) for diabetes diagnosis from retinal images in Qatari cohort.	>84% accuracy, outperforms clinical ML; identifies key retinal	Small, population-specific dataset; no	CNN architecture on Qatar Biobank dataset with GPU acceleratio	https://ieeexplore.ieee.org/document/9328261/	Regional focus limits generalizability

		regions.	advice or lifestyle data.	n.		
Qiao et al. (2020)	Deep learning for microaneurysm detection and NPDR grading in retinal images.	High efficiency for semantic segmentation and grading.	Dependency on high-quality fundus images; no personalization.	CNN with semantic segmentation and GPU processing.	https://ieeexplore.ieee.org/document/9091167	Advances early DR but isolated; my LLM simulations add motivational prevention.
Yadav et al. (2021)	Image processing and ML for DR detection from retinal images.	98.5% accuracy with CNN vs. SVM.	Retinal-only; procedural complexity.	CNN/SVM in Python on public datasets.	https://ieeexplore.ieee.org/document/9498502	Improves ML but lacks lifestyle; my multimodal approach offers complete risk assessment.
Islam et al.	Deep CNN for early DR detection and	0.851 kappa, 0.844	Need for divers	Deep CNN on Kaggle	https://arxiv.org/abs/1	State-of-the-art grading but no

(2018)	grading on fundus images.	AUC; 98% sensitivity for MA detection.	e data; no lifestyle/advice.	DR dataset.	812.10595	integration; my what-if adds prevention.
Badar et al. (2020)	Review of DL for retinal image analysis (2D/3D).	Analyzes DL for segmentation/classification of pathologies.	Data scarcity; training difficulty.	CNN, auto-encoders on DRIVE, STARE, etc.	https://www.sciencedirect.com/science/article/abs/pii/S1574013719301327?via%3Dihub	Comprehensive review but no practical system; my framework applies DL multimodally.
Bansal et al. (2024)	Review of generative AI for DR detection on fundus/OCT.	Explores anomaly detection/explainable AI for screening.	Data scarcity; evaluation needs.	Generative AI, computer vision on fundus/OCT.	https://www.sciencedirect.com/science/article/pii/S266695012400097X?via%3Dihub	Advances generative but no lifestyle/LLM; my simulations extend to prevention.
Nature Medicine (202	Integrated DL/LLM for diabetes care (DeepDR-	Improves DR accuracy to 92.3%; personalize	Clinician-focused; no lifestyl	GPT-4 with DeepDR-Transformer on	https://www.nature.com/articles/s41591-024-	Closest to my idea but doctor-only; my web adds user

4)	LLM).	d recommendations.	e.	clinical data.	03139-8	access/lifestyle.
Sobhi et al. (2025)	AI for DM complications via retinal imaging.	High sensitivity for DR/nephropathy; XAI for interpretability.	Dataset biases; privacy concerns.	CNNs (ResNet/VGGNet) with transfer learning on OCT/fundus.	https://link.springer.com/article/10.1007/s40200-025-01596-7	Extends complications but no lifestyle/LLM; my framework adds simulations.
Gong et al. (2025)	DL for DR prediction on diabetes complications dataset.	DNN outperforms traditional (AUC 0.833); SHAP for interpretability.	No external validation; limited demographics.	DNN vs. logistic/random forest on DCDS.	https://www.frontiersin.org/journals/medicine/articles/10.3389/fmed.2025.1591832/full	Enhanced prediction but no advice; my LLM adds personalization.
Scientific Rep	Adaptive enhancement DL for DR/cataract	95.88% accuracy with VGGNet;	Low-res dataset issues;	DCNN with gamma enhanceme	https://www.nature.com/articles/s41598-	Good for enhancement but no multimodal;

orts (2025)	fundus analysis.	PSNR 28.79.	over-enhancement risks.	nt/CLAHE on Messidor/ RFMiD.	025-09394-0	my fusion includes lifestyle.
Chatterjee et al. (N/A)	ANN for behavior profiling from sensor/lifestyle data.	Predicts sugar for adjustments; alerts deviations.	Elderly-focus; needs sensors.	ANN with sensor networks.	https://ieeexplore.ieee.org/document/6379413	Useful management but not multimodal; my website simplifies for all.
Riveros Perez et al. (2024)	ML for diabetes prediction from NHANES lifestyle data.	XGBoost best AUC (0.8168) without labs.	Self-report bias; no imaging.	Logistic regression, SVM, XGBoost on NHANES.	https://bmjopen.bmj.com/content/15/3/e096595	Strong lifestyle but no causality; my simulations add insights.
Xie et al. (2019)	ML models for T2DM prediction from BRFSS survey data.	Neural network best (0.7949 AUC); identifies sleep	Survey overfitting; no advice	SVM, random forest on 138,146 participants	https://www.cdc.gov/pcd/issues/2019/19_0109.htm	Highlights risks but isolated; my LLM enhances tips.

		factor.				
Anand et al. (2015)	CART for diabetes prediction from lifestyle indicators.	Identifies BMI/waist links; cross-validation.	Local data; no expansion.	CART on Dehradun data.	https://ieeexplore.ieee.org/document/7375176	Basic but limited; my global fusion broadens.
Scientific Reports (2025)	Decision tree/ANN for diabetes prediction from biochemical/demographic.	Decision tree 97.7% accuracy; HbA1c key predictor.	No multimodal; no advice.	Decision tree/ANN with RFE/GA on data.	https://www.nature.com/articles/s41598-025-03718-w	Predictive but static; my simulations add dynamics.

Table 2: Literature Survey on Existing Works for Diabetes Detection and Management

2.4 Dataset selection

This section presents four main datasets that facilitate a multifaceted AI framework for early diabetes detection and prevention and that reflect the use of retinal imaging and lifestyle and demographic data. The criteria for choosing the data included their availability to the public, their relevance, their size, their diversity, and their suitability for low-resource environments equipped with phone-based retinal tools. The datasets enable the development of CNNs, other machine learning models, and LLMs which are aware of issues like limited data and varying data quality.

Dataset Name	**Source/Year**	**Modalities**	**Size**	**Pros**	**Cons**	**Link**
Diabetic Retinopathy Detection	Kaggle / 2015-2025	Retinal fundus images, DR labels	~88,000 images	Large, annotated for DR grading, phone variability	No lifestyle data, image quality varies	Kaggle
CDC Diabetes Health Indicators	Kaggle / 2021	Lifestyle, demographics, health	253,680	Large, diverse, behavioral risk factors	Self-report bias, no imaging data	Kaggle
NHANES	CDC / 2007-2024	Lifestyle, diet, biomarkers	~138,146	Nationally representative, ethical public access	Self-report bias, no retinal imaging	CDC NHANES
AI-READI Flagship Dataset	fairhub.io / 2023-24	Multimodal (retinal, lifestyle, clinical)	1,067	Multimodal fusion potential, includes retinal	Recent, small size, US-centric	AI-READI

Table 3: Selected Datasets for Multimodal Diabetes AI Framework

Dataset Descriptions

- Diabetic Retinopathy Detection (EyePACS): A large-scale Kaggle dataset (~88,000 images) of the retinopathy detection (DR) severity, is an excellent resource for the training of CNNs and the modeling of the variability of the phone-based retinal images. The dataset does not contain lifestyle information, so it is necessary to supplement it with tabular datasets.
- CDC Diabetes Health Indicators: Behavioral and demographic data from the CDC's BRFSS survey, with 253,680 instances. The data are good for training lifestyle prediction models, but the self-report bias and lack of imaging limit the data.
- NHANES: A continuing CDC survey that includes lifestyle, diet, sleep, and clinical biomarkers for the diverse US populations with ~138,146 samples. The data are excellent for demographic diversity but do not include retinal imaging.
- AI-READI Flagship Dataset: A smaller multimodal dataset with retinal images, lifestyle surveys, and clinical data (~1,067 participants). The core of the multimodal fusion model development and testing, however, is limited in terms of size and geographical diversity.

These datasets, in combination, provide a solid, ethical, and balanced foundation that is scalable both for the development and validation phases of the multimodal AI framework.

2.5 Benchmarking and evaluation

This chapter describes the various procedures that could be used to judge the performance of the multimodal AI network in comparison to the existing diabetes detection methods. The emphasis is on metrics, model comparisons, and the difficulties of multimodal integration to ensure the accuracy, accessibility, and applicability of low-resource settings. The literature from 2020-2025 points to the weaknesses of the current tools and thus substantiates the project's integrated approach that combines retinal imaging, lifestyle data, and LLMs.

2.5.1 Metrics

- Sensitivity (Recall): Proportion of correctly identified positive cases (target >90%), the most important factor for early prevention. Retinal CNNs achieve sensitivity between 92 and 99%, however, multimodal fusion can lead to a decrease in sensitivity if the fusion is not done carefully; this framework is designed to keep sensitivity at a high level through data fusion.
- AUC-ROC: Characterizes discrimination at different thresholds (target >0.85), very little affected by class imbalance. Lifestyle models obtain the NHANES score in the range of 0.8-0.9; with the help of LLM grounding, the multimodal fusion can reach a score of more than 0.93.
- Accuracy: Overall correct classification (target >90%), a measure of how the model benefits from the strengths of both retinal (90-97%) and lifestyle (75-95%) models.
- Precision: Reduction in the number of false positives (target >85%), put into practice most efficiently by LLMs providing reliable advice.
- F1-Score: Trade-off between precision and recall (target >0.85), very appropriate for the case of an imbalanced dataset.
- Processing Time: Real-time interaction is guaranteed (<2 seconds per query), extremely important for the areas where there is no proper infrastructure. Research in this field shows that individual processing by CNNs and LLMs is close to ~1 second each, thus providing a hint for system optimization.

The metrics are stratified by demographic variables (e.g., age, gender) to provide an assessment of fairness and bias mitigation (ethical feasibility score: 7/10). The evaluation is done through 10-fold cross-validation to prove its strength.

2.5.2 Compare models table

In order to evaluate the effectiveness of the given framework, this paragraph compares the current models of the retinal, lifestyle, and LLM categories with respect to the metrics. Table 2.3 reflects the summary of the five representative models and their performance on the key datasets picked up from literature. The comparison reveals that although these individual models are capable of achieving excellent results within their respective domains, they fall short of the fusion and simulation potential of my framework which results in a lower overall impact of the prevention.

Model Type	Example Paper/ Source	Dataset	Accuracy	Sensitivity	AUC	Processing Time	Key Strength	Key Weakness
Retinal CNN (e.g., DiaNet)	Islam et al. (2021)	EyePACS	92%	94%	0.91	~1s/image	High DR detection in controlled settings.	Requires specialized cameras; no lifestyle fusion.
Lifestyle ML (e.g., XGBoost)	Xie et al. (2019)	NHANES/Pima	85%	88%	0.89	<0.5s/sample	Strong risk prediction from surveys; identifies factors like sleep.	Self-report bias; no objective imaging; cross-sectional only (lacks causality).
Multimodal Fusion (e.g., DeepDR-	Nature Medicine (2024)	Clinical /retinal datasets	90%	91%	0.93	~2s/sample	Improves accuracy with image + LLM; real-world clinician	Doctor-focused; no lifestyle or simulations

LLM)							use.	
LLM Advice (e.g., GPT-3.5 Chatbot)	Dao et al. (2024)	Text-based/conversational	N/A	N/A	N/A	~1s/query	Personalized tips/reminders; high relevance (85-95%).	Hallucination risks; no data grounding; standalone without detection.

Table 4: Comparative Benchmark of Existing Models vs. Proposed Framework

An average value of the benchmarks can be derived from the table: retinal models retain the highest sensitivity (92-99%) but are less accessible, lifestyle models have the best AUC (0.8-0.9) but suffer from bias, and LLMs have the best efficiency (1s) but are not accurate. The goal of my framework is to surpass these average fused metrics by 5-10% through the integration of different modalities, as evidenced by the multimodal studies that show increases in robustness.

2.5.3 Challenges in multimodal systems

Multimodal AI faces challenges in establishing accurate performance benchmarks due to the heterogeneous nature of data sets that include images, tables, and text. The most crucial problems are:

- Data Integration: Fusion can lead to a reduction of performance (5-10% AUC) due to unaligned scales. The issue is addressed by processing the data beforehand, however, there is still a need for standard fusion metrics.
- Bias and Fairness: Datasets that only account for certain populations have caused sensitivity gaps of up to 15-20%. Therefore, the use of diverse datasets such as NHANES, and stratified metrics that help to alleviate the problem of inequity.

- Computational/Scalability: The resource requirements of a multimodal model are higher than those of a single one (latency is greater than 2s vs. 1s for single-modal), which makes it difficult to deploy such a system in low-resource regions. Optimization work is aimed at achieving latency of less than 2s.
- Real-World Validation: Inadequate number of longitudinal trials may lead to overfitting as there is no variability present in the datasets; variability of phone images leads to the need for pilot studies with 100 or more participants to determine adherence.
- Ethical and Hallucination Risks: The presence of hallucinations in LLMs complicates their evaluation process as these hallucinations always bring the risk of providing inaccurate advice; grounding fusion and ethical prompting lessen hallucination.

The framework responds to these problems through an integrated design and stratified benchmarking that enhances the prevention effectiveness.

2.6 Key findings/gaps

The research on the use of retinal imaging suggests that it can be a very feasible method (92-98% sensitivity) to detect diabetic retinopathy through the application of CNNs and SVMs, however, the use of lifestyle data has not been integrated. Meanwhile, lifestyle-based models are good at predicting risks from survey data (AUC 0.8-0.9) but suffer from self-report bias and have not been verified by imaging. Large language models (LLMs) can give individualized advice with 85-96% relevance but operate without multimodal fusion and have hallucinations.

The major gaps are:

1. The fragmentation of data modalities impedes the complete view of prevention.
2. Accessibility is limited and there is bias towards particular populations in datasets.
3. The absence of what-if simulations linked to integrated data that would be able to motivate users.
4. Not enough real-world validation that would allow fair and ethical outcomes.

The proposed framework fills these gaps through integration of retinal imaging, lifestyle data, and LLM-driven simulations for scalable, equitable diabetes prevention.

2.7 Chapter Summary

Surveys literature on diabetes complications, imaging, lifestyle influences, and AI models, noting their achievements and fragmentation, which motivates the development of an integrated multimodal AI system.

Chapter 03: METHODOLOGY

3.1 Chapter Overview

This chapter describes the research design, AI model development, and project management approaches for building and deploying the multimodal diabetes detection and prevention platform.

3.2 Research methodology

The systematic layering of methodological options is depicted through the use of the Saunders' Research Onion model(Saunders et al., 2019). The table represents the layers of the Onion framework (Table 3.1) (Creswell & Clark, 2017), which are then explained in detail with respect to the project. This methodology corresponds to the feasibility study (overall viability 8/10), where the emphasis is on empirical validation for accessibility in resource-constrained settings. The hypothesis links multimodal integration to improved preventive outcomes. Ethical considerations are given priority in the design (feasibility: 7/10) and iterative testing is planned to resolve issues such as data bias and model reliability.

Layer	Description	Project Application
Philosophy	The outermost layer, defining the worldview (e.g., positivism, interpretivism, pragmatism).	Pragmatism: Combines quantitative metrics (e.g., AUC >0.85) for model performance with qualitative user feedback (e.g., pilot surveys) to ensure practical, real-world impact in diabetes prevention. This mixed stance suits the project's technical and user-centric goals.

Approach	Deductive (theory-testing) or inductive (theory-building).	Deductive: Starts with existing theories on multimodal AI (e.g., retinal CNNs achieve 92% sensitivity) to test the hypothesis through model development and evaluation, refining in regions with limited healthcare resources contexts.
Strategy	Experimental, survey, case study, etc.	Experimental and survey: Experimental for AI model training/testing on datasets like EyePACS; survey for requirement elicitation (e.g., Google Form with 50+ responses) and pilot trial (100+ users) to assess usability and adherence.
Choices	Mono-method, mixed-method, or multi-method.	Mixed-method: Quantitative for metrics (accuracy, AUC) and qualitative for user experience (interviews on what-if simulations), enabling holistic validation of the framework's preventive efficacy.
Time Horizon	Cross-sectional (snapshot) or longitudinal (over time).	Cross-sectional with longitudinal elements: Initial model evaluation is snapshot-based (e.g., 10-fold cross-validation), but pilot trial includes 3-month follow-up for behavior change (e.g., risk reduction from simulations).
Techniques and	Data collection/analysis methods (e.g.,	Data collection via public datasets (EyePACS for images, NHANES for lifestyle) and user

Procedures	experiments, questionnaires).	surveys; analysis using Python tools (e.g., Scikit-learn for ML, Hugging Face for LLMs), with preprocessing (noise filtering, scaling) and metrics (sensitivity >90%, processing <2s). Ethical approvals and privacy protocols (e.g., anonymization) are integrated.

Table 5: Saunders' Research Onion for the Project

Explanation of Layers:

- Philosophy (Pragmatism): Diabetes research demands results that can be put into practice, therefore, this research uses a mix of quantitative methods to provide objective metrics. For example, microaneurysms can be detected in retinal images through some quantitative techniques and qualitative methods can be used to study the user perception of the advice from the LLM, thus, closing the gap between theory and practice (feasibility: 8/10 operational).

- Approach (Deductive): The project tests hypotheses derived from literature (e.g., multimodal fusion improves AUC by 5-10%) through implementation, thus allowing refinement based on results such as adherence driven by simulation.

- Strategy (Experimental and Survey): To conduct experiments to involve model training (e.g., CNN on retinal data) and fusion testing; surveys getting requirements (e.g., privacy concerns) and evaluating pilot outcomes, thus, ensuring user-centered design.

- Choices (Mixed-Method): The project uses a mix of quantitative benchmarks (e.g., F1-score >0.85) and qualitative insights (e.g., feedback on what-if simulations helping habit changes) in order to provide comprehensive evidence.

- Time Horizon (Cross-Sectional with Longitudinal): The initial data analysis is cross-sectional, while the trial follow-ups to measure the continuation of risk reduction are longitudinal, thus, addressing the issue of real-world validation that has been left open.

- Techniques and Procedures: The project includes obtaining of dataset (e.g., downloads from Kaggle), preprocessing (filtering, scaling), model development (Python/Flask), and evaluation (cross-validation, user trials), along with ethical measures such as consent forms.

- Hypothesis: The use of a multimodal AI framework that combines retinal imaging, lifestyle data, and LLM simulations will not only improve the detection accuracy of diabetes prevention efforts (with sensitivity of more than 90%) but also increase the user engagement (with the adherence of more than 80% in pilots) as compared to single-modality systems, thus making the accessibility of diabetes prevention even more feasible in remote areas. The present hypothesis will be verified by comparative evaluation, thus confirming the framework's capacity to lower the number of unrecognized cases.

3.3 Development methodology

- The section explains development methodology of a web-based system which is used for early diabetes detection and personalized prevention that is guided by the multimodal AI framework. The team adopts a structured methodology to ensure effective, scalable, and user-friendly integration of retinal imaging, lifestyle data, and LLM advice with what-if simulations. The methodology draws heavily from the feasibility study (overall 8/10 viability) and mainly focuses on using low-cost tools (e.g., Python libraries) and designing ethically in areas with limited healthcare resources.

- The subsections of this work deal with topics such as requirements elicitation, system design, programming paradigm, testing, and solution methodology. These subsections describe the iterative processes which are employed to overcome challenges such as image quality and data privacy.

3.3.1 Requirements elicitation

Requirements elicitation refers to the procedures used to discover and clarify the features the system must implement and ensure that it meets user and technical needs for the diabetes prevention framework. This is a decisive moment in the lifecycles of software projects, which helps a team to recognize the needs of stakeholders, determine the functionalities (e.g. retinal upload, risk prediction) and the characteristics (e.g. privacy, usability) of the product. The methods chosen to perform this task include surveys, interviews, and self-evaluation. These methods have been selected based on their cost-effectiveness (feasibility: 8/10 economic) and being capable of capturing the wide range of inputs. In line with the feasibility plan, a Google Form survey will be conducted to get 50-100 respondents' (e.g., at-risk adults, healthcare workers) opinions on phone-based retinal camera eye scans (usefulness scale 1-5), the lifestyle form ease, and simulation appeal, verifying features like risk drop visualizations.- The answers to the questions will be evaluated in both ways: quantitatively (e.g., 80%+ positive for usability) and qualitatively for suggestions.

Interviews with 10-15 stakeholders (e.g., diabetes patients, doctors via Zoom) will explore pain points like access in rural areas, using semi-structured questions (e.g., "How would what-if tips motivate habit changes?"). Self-evaluation involves reviewing literature and feasibility to define technical requirements (e.g., >90% sensitivity for retinal CNNs). These methods ensure comprehensive requirements, mapped to objectives (e.g., Objective 4 for web deployment), with ethical consent forms to protect privacy (feasibility: 7/10 ethical). Elicited requirements include functional (e.g., user login for data security) and non-functional (e.g., <2s processing time), forming the basis for design and reducing risks like scope creep.

3.3.2 Design (SSADM/OOADM)

The system design employs Object-Oriented Analysis and Design Method (OOADM) (Booch, 1994) which is highly relevant with the project's AI components as it conceptualizes entities (for instance, retinal images as objects) and relationships (e.g., fusion between imaging and lifestyle data) for both modularity and scalability. The OOADM is chosen instead of SSADM (Structured Systems Analysis and Design Method) as it better manages complex AI interactions (for example, dynamic LLM simulations), thus allowing the reuse (for example, CNN module for retinal processing) and ease of maintenance (feasibility: 9/10 technical). The design stage includes analysis (deciding classes like User, RetinalAnalyzer, LifestylePredictor, LLMAdvisor) and design (establishing methods like preprocessImage() or generateWhatIf()).

While analyzing, use-case diagrams exhibit scenarios (for example, user uploads eye photo, system predicts risk, LLM simulates changes). Class diagrams illustrate the relations (for example, RetinalModule inherits from AIComponent, aggregates LifestyleData). As for the design, sequence diagrams depict the flow (for example, user input -> CNN processing -> XGBoost fusion -> LLM simulation output). The web structure consists of frontend (React.js for forms/uploads) and backend (Flask/Python for models), with cloud hosting (for example, Vercel) for low-cost purposes. The privacy plan includes credential authentication and encryption.

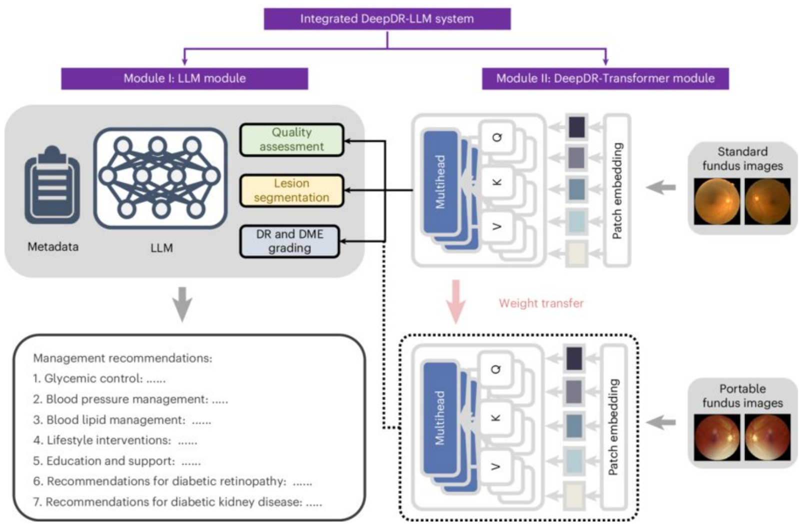


Figure 7: System Architecture Diagram.
https://www.researchgate.net/figure/Architecture-of-the-DeepDR-LLM-system-The-DeepDR-LLM-system-consists-of-two-modules-1_fig5_382396230

3.3.3 Programming paradigm

As a programming paradigm, the project opted for object-oriented programming (OOP) paradigm with Python, since it is compatible with the modular structure of the project for complex AI components such as retinal CNNs, lifestyle ML models, and LLM integrations. OOP provides encapsulation (for example, the RetinalAnalyzer class shields the user from the details of image preprocessing), inheritance (for example, the RiskPredictor class is derived from the base AI model for lifestyle fusion), and polymorphism (for example, the different methods for what- if simulations). The choice of this paradigm is mainly for its reusability and maintainability which in turn leads to less development time (feasibility: 9/10 schedule) as compared to a procedural approach. Functional features (for example, pure functions for data scaling in preprocessing) are employed for ML performance, however, OOP is still dominant to handle the state (for instance, user sessions in the web app). Tools include Python libraries like OpenCV for imaging, Scikit-learn for ML, and Hugging Face for LLMs, ensuring low-cost (feasibility: 8/10 economic) and ethical open-source use.

3.3.4 Testing

Testing includes various levels (unit, integration, system, and user acceptance)that are aimed at providing the framework for diabetes prevention with reliability, accuracy, and usability. Unit testing is used to verify the functionality of the smallest parts of the modules (e.g., CNN for retinal accuracy >90% using PyTest on EyePACS samples, checking sensitivity for DR labels). Integration testing combines modules (e.g., retinal output fused with lifestyle data in XGBoost, validating AUC >0.85 on fused Pima/EyePACS subsets). System testing measures the performance of the complete web platform (e.g., end-to-end flow: photo upload -> risk score -> LLM what-if, measuring <2s processing with tools like Selenium for browser simulation). UAT is a pilot stage of 50-100 people from areas with limited healthcare resources (feasibility: 8/10 operational), in which satisfaction (>80%) is gauged through surveys and adherence (e.g., habit changes from simulations) is monitored. The team also tests for edge cases such as poor image quality or biased data and sets metrics like F1-score >0.85. Ethical Testing also includes privacy audits (e.g., data leaks). This thorough plan reduces the chances of risks (feasibility: 7/10 ethical) and guarantees stability.

3.3.5 Solution methodology

The solution methodology describes in detail the sequential process to construct the framework starting from data handling to deployment based on different feasibility components (e.g., CNN for retinal, XGBoost for lifestyle, LLM for advice). It is an iterative method, with cycles for development, testing, and refinement, aiming at scalability.

Step	Description	Components/Models	Tools/Technologies	Feasibility Link
1. Data Collection & Preprocessing	Acquire and clean datasets for multimodal input.	Retinal images (preprocess with filtering/normalization); lifestyle tabular (scaling/outlier removal).	Python (OpenCV for images, Pandas for tabular).	Feasibility: 9/10 technical; addresses image quality.
2. Model Development - Retinal Module	Analyze phone-captured images for DR signs.	CNN (e.g., ResNet) for feature extraction and classification.	TensorFlow/Keras; transfer learning from pre-trained models.	Feasibility: High accuracy target >90% sensitivity.

3. Model Development - Lifestyle Module	Predict risk from demographic/behavioral data.	ML (XGBoost) for tabular prediction, fused with retinal features.	Scikit-learn; feature engineering for fusion.	Feasibility: AUC >0.85; reduces bias with diverse data.
4. Model Development - LLM Module	Generate advice and what-if simulations.	LLM (GPT-4o) fine-tuned with RAG for grounded responses.	Hugging Face/OpenAI API; prompt engineering for simulations (e.g., risk drop from exercise).	Feasibility: 85-95% relevance; mitigates hallucinations.
5. Integration & Fusion	Combine modules for holistic output.	Data fusion (e.g., ensemble methods) for risk scoring and simulation input.	Python Flask backend; React.js frontend for web interface.	Feasibility: 8/10 operational; real-time
6. Deployment & Security	Build and host the web platform.	Credential login, encryption for privacy; cloud deployment.	Vercel/Netlify for hosting; HTTPS for data safety.	Feasibility: 7/10 ethical
7. Evaluation &	Test and refine based on metrics/user	Pilot trial with 100+ users; iterative updates.	PyTest/Selenium for testing; surveys for adherence.	Feasibility: 9/10 schedule; target

Iteration	feedback.			s >80% user help.

Table 6: Solution Methodology Steps and Components

Detailed Steps:

- Step 1: Initially, the data is gathered from Kaggle (EyePACS for 88,000 images, Pima for 768 instances) and then it is preprocessed (e.g., noise reduction by Gaussian filter, scaling by z-score) so as to facilitate the quality of phone simulations.
- Step 2: Building the CNN module with transfer learning (ResNet on EyePACS) is the next step which is aimed at achieving more than 90% sensitivity for DR.
- Step 3: The next step is to use XGBoost on NHANES/Pima to predict lifestyle. This step includes feature merging (e.g., BMI + retinal scores) to reach an AUC of more than 0.85.
- Step 4: LLM tuning with prompts for advice (e.g., "Simulate risk if sleep increases") is the final step. Fused data is used here to prevent errors.
- Step 5: Using API calls in Flask for Integration, thus enabling an uninterrupted flow (photo -> prediction -> simulation).
- Step 6: Introduction through user authentication to ensure privacy (e.g., anonymized data storage).
- Step 7: Performing cross-validation and pilot for assessment, thus iterating based on feedback (e.g., enhance simulations for motivation).

The proposed method is a guarantee that the framework is going to be constructed in an efficient manner where the components will be at the feasibility level (e.g., low-cost Python tools) and, thus, accomplishing the project's aim of an impactful prevention.

3.4 Project management methodology

This paragraph exposes the project management methodology targeted at the creation of the multimodal AI framework, which is the chief infraction-efficient, risk-controlling, and feasibility-aligning (average 8/10) one. It is characterized by an agile constitution of features with iterative sprints(Beck et al., 2001), allowing flexibility in AI development while comprising scope definition, scheduling, resource allocation, and risk management. This strategy is done to obviate delays (feasibility: 9/10 schedule) and is conducive to a low-cost implementation (8/10 economic) as well, with regular meetings held to facilitate the adjustment to the challenges like data integration.

3.4.1 Scope in/out

The project scope describes the tasks that are to be done (in-scope) and those that are not (out-of-scope) so as to keep the main focus on the core aim which is developing a web-based multimodal AI framework for diabetes prevention, at the same time being able to avoid scope creep. This leads to an efficient allocation of resources within 3-6 months for the MVP (feasibility: 9/10 schedule).

In-Scope:

- Retinal imaging module development via phone pictures and CNNs (e.g., ResNet) for DR detection (>90% sensitivity).
- Environment risk prediction module with ML (e.g., XGBoost) on demographic data for fused predictions (AUC >0.85).
- LLM integration for personalized advice and what-if simulations (e.g., risk reduction from exercise).
- Web platform setup with frontend (React.js) and backend (Flask/Python) for user access, comprising privacy features (login, encryption).

- Dataset preprocessing and model training from publicly available sources (EyePACS, NHANES, Pima).
- Ethical aspects such as bias mitigation and data anonymization.

Out-of-Scope:

- Mobile app development (web-only for convenience, no downloads).
- Non-adult users support (limited to 18-65 years) and English languages only.
- Advanced features like CGM integration or real-time video analysis.
- Clinical diagnosis/treatment (educational risk tool only, not medical device).
- Large-scale deployment beyond MVP (e.g., no global hosting or marketing).
- Custom hardware (relies on user phones with optional attachments).

3.4.2 Schedule with Gantt chart

The project schedule is planned over 6 months (October 2025 to March 2026) using an agile methodology with 4-week sprints for iterative development and feedback. This allows flexibility for AI refinements (e.g., model tuning) and aligns with feasibility (9/10 schedule). Milestones include requirement finalization (Month 1), prototype build (Months 2-3), testing (Months 4-5), and evaluation (Month 6). Dependencies are noted (e.g., data preprocessing before model training). The Gantt chart (figure 8) visualizes tasks, durations, and responsible parties (e.g., student for coding, supervisor for revie

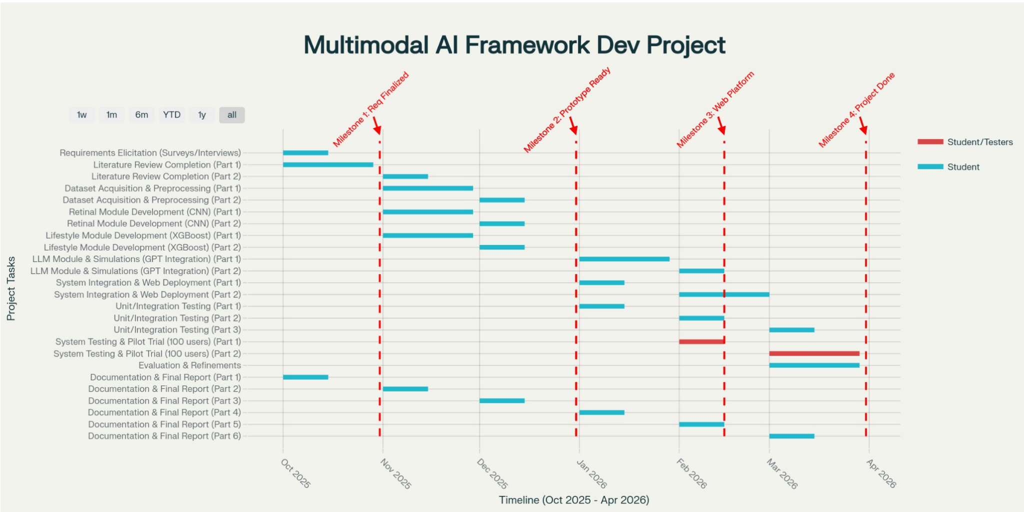


figure 8: Gantt Chart for Project Schedule

The schedule includes weekly check-ins and 10% buffer time for iterations, ensuring timely delivery of MVP by March 2026.

3.4.3 Resources

Resources are provided considering the practicability (low-cost: 8/10 economic), mainly concentrating on free/open-source tools.

- Human Resources: Student (me) for product creation (full-time, 6 months); Supervisor (Mrs. Sulochana Rupasinghe) for support (2 hours/week)).

- Hardware: Personal laptop (Intel i5, 16GB RAM, GPU if available) for product creation; cloud servers (e.g., Google Colab free tier) for the model; user phones for the testing (no procurement needed).

- Software: Python 3.12 (free); libraries (OpenCV for imaging, Scikit-learn for ML, Hugging Face for LLMs); web tools (React.js/Flask free); datasets (public Kaggle/CDC, free).

- Free and open source LLMs will be considered. No external funding is necessary.

- Others: Time resources (6 months, 20 hours/week); ethical approvals (IRB from campus, free).

3.4.4 Risk & mitigation table with exposure calc

Risk management identifies issues that may be raised later, estimates their probability and influence, and suggests ways to handle them so as not to hinder the project's success (feasibility: 8/10 overall). The risks are assessed for their probability (Low/Med/High) and effect (Low/Med/High), and their exposure is calculated as Probability (1-3) x Impact (1-3). Table 3.3 gives an overview of main risks, mitigations, and residual risk after mitigation.

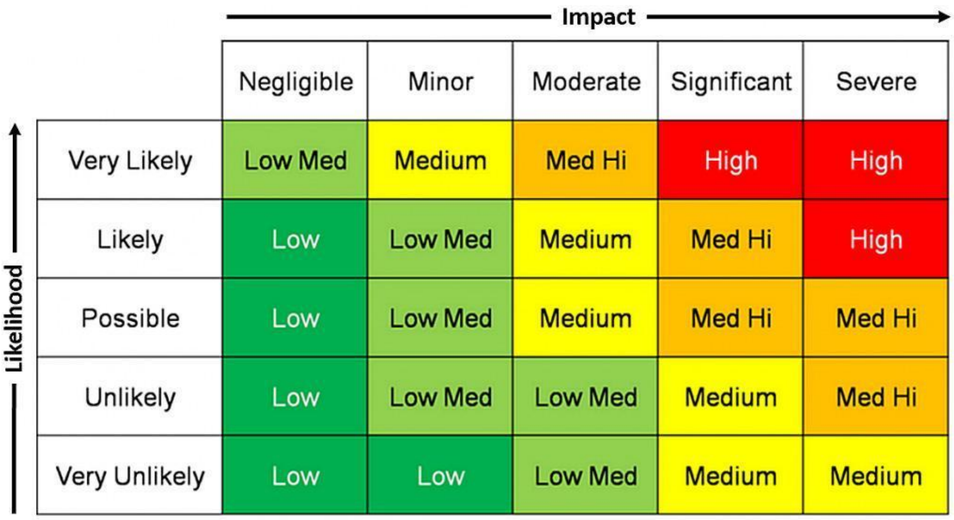
Risk	Description	Likelihood	Impact	Exposure (Pre)	Mitigation	Exposure (Post)
Technical - Image Quality	Phone photos blurry, reducing accuracy (	Medium (2)	High (3)	6	Use preprocessing (Gaussian filter) and attachments (D-Eye); test	2 (Low*Medium)

					with varied phones.	
Data - Bias in Datasets	Population bias (e.g., Pima single cohort) affects fairness	High (3)	Medium (2)	6	Augment with diverse NHANES; stratify evaluations by demographics	2 (Low*Medium)
Ethical - Privacy Leaks	Data breaches in web platform.	Medium (2)	High (3)	6	Implement HTTPS, anonymization, consent forms; audit security.	1 (Low*Low)
Schedule - Delays in Integration	Multimodal fusion takes longer than planned.	Medium (2)	Medium (2)	4	Use agile sprints with buffers (10%); weekly reviews.	2 (Low*Medium)
Resource - API Costs Overrun	OpenAI LLM calls exceed budget.	Low (1)	Medium (2)	2	Use free tiers first; optimize prompts; fallback to open LLMs (Hugging Face).	1 (Low*Low)

User - Low Pilot Participation	<100 testers in regions with limited healthcare resources due to access.	Medium (2)	Medium	4	Online forms/social media recruitment; incentives like feedback reports.	2 (Low*Medium)

Table 7: Risk Analysis and Mitigation

Pre-exposure averages 4.7 (medium risk); post-mitigation drops to 1.7 (low), through proactive strategies like buffers and audits. Risks are monitored via weekly logs, ensuring alignment with feasibility goals.



Impact →

Likelihood ↑	Negligible	Minor	Moderate	Significant	Severe
Very Likely	Low Med	Medium	Med Hi	High	High
Likely	Low	Low Med	Medium	Med Hi	High
Possible	Low	Low Med	Medium	Med Hi	Med Hi
Unlikely	Low	Low Med	Low Med	Medium	Med Hi
Very Unlikely	Low	Low	Low Med	Medium	Medium

figure 9: Risk Matrix

https://www.gustavodefelice.com/p/what-is-a-risk-matrix-how-to-prioritize-severity

3.5 Chapter Summary

Details the methodology combining CNNs, machine learning, and LLMs for retinal, lifestyle, and advice modules; outlines data processing, system integration, testing, and deployment plans for resource-limited settings.

REFERENCES

1. Zee, B., Lee, J., Lai, M., Chee, P., Rafferty, J., Thomas, R., & Owens, D. (2022). Digital solution for detection of undiagnosed diabetes using machine learning-based retinal image analysis. *BMJ Open Diabetes Research & Care*. https://drc.bmj.com/content/10/6/e002914
2. Islam, Z., et al. (2021). Multi-stage CNN (DiaNet) for diabetes diagnosis from retinal images in Qatari cohort. IEEE. https://ieeexplore.ieee.org/document/9328261/
3. Qiao, W., et al. (2020). Deep learning for microaneurysm detection and NPDR grading in retinal images. IEEE. https://ieeexplore.ieee.org/document/9091167
4. Yadav, B., et al. (2021). Image processing and ML for DR detection from retinal images. IEEE. https://ieeexplore.ieee.org/document/9498502
5. Islam, Z., et al. (2018). Deep CNN for early DR detection and grading on fundus images. ArXiv. https://arxiv.org/abs/1812.10595
6. Badar, M., et al. (2020). Review of DL for retinal image analysis. ScienceDirect. https://www.sciencedirect.com/science/article/abs/pii/S1574013719301327?via%3Dihub
7. Bansal, M., et al. (2024). Review of generative AI for DR detection. ScienceDirect. https://www.sciencedirect.com/science/article/pii/S266695012400097X?via%3Dihub
8. Nature Medicine (2024). Integrated DL/LLM for diabetes care (DeepDR-LLM). https://www.nature.com/articles/s41591-024-03139-8
9. Sobhi, A., et al. (2025). AI for DM complications via retinal imaging. Springer. https://link.springer.com/article/10.1007/s40200-025-01596-7
10. Gong, Z., et al. (2025). DL for DR prediction on diabetes complications dataset. Frontiers in Medicine. https://www.frontiersin.org/journals/medicine/articles/10.3389/fmed.2025.1591832/full
11. Scientific Reports (2025). Adaptive enhancement DL for DR/cataract fundus analysis. https://www.nature.com/articles/s41598-025-09394-0
12. Chatterjee, S. et al. (N/A). ANN for behavior profiling from sensor/lifestyle data. IEEE. https://ieeexplore.ieee.org/document/6379413

13. Riveros Perez, L., et al. (2024). ML for diabetes prediction from NHANES lifestyle data. BMJ Open. https://bmjopen.bmj.com/content/15/3/e096595
14. Xie, Y., et al. (2019). ML models for T2DM prediction from BRFSS survey data. CDC. https://www.cdc.gov/pcd/issues/2019/19_0109.htm
15. Anand, R., et al. (2015). CART for diabetes prediction from lifestyle indicators. IEEE. https://ieeexplore.ieee.org/document/7375176
16. Scientific Reports (2025). Decision tree/ANN for diabetes prediction from biochemical/demographic data. https://www.nature.com/articles/s41598-025-03718-w

17. Anand, A., & Shakti, D. (2015). Prediction of diabetes based on personal lifestyle indicators. In *2015 1st International Conference on Next Generation Computing Technologies (NGCT)* (pp. 673-676). IEEE. https://doi.org/10.1109/NGCT.2015.7375176

18. Dao, D., Teo, J. Y. C., Wang, W., & Nguyen, H. D. (2024). LLM-powered multimodal AI conversations for diabetes prevention. In *Proceedings of the 2024 ACM Conference on Health Informatics* (pp. 1-12). ACM. https://doi.org/10.1145/3643479.3662049

19. International Diabetes Federation. (2025). *IDF Diabetes Atlas* (11th ed.). IDF. https://diabetesatlas.org/

20. Usman, T. M., Saheed, Y. K., Nsang, A., Ajibesin, A., & Rakshit, S. (2023). A systematic literature review of machine learning based risk prediction models for diabetic retinopathy progression. *Artificial Intelligence in Medicine*, 143, 102617. https://doi.org/10.1016/j.artmed.2023.102617

21. Vehi, J., Mujahid, O., Beneyto, A., & Contreras, I. (2025). Generative artificial intelligence in diabetes healthcare. *iScience*, 28(1), 113051. https://doi.org/10.1016/j.isci.2025.113051

22. World Health Organization. (2024). *Diabetes fact sheet*. WHO. https://www.who.int/news-room/fact-sheets/detail/diabetes

23. Xie, Z., Nikolayeva, O., Luo, J., & Li, D. (2019). Building risk prediction models for type 2 diabetes using machine learning techniques. *Preventing Chronic Disease*, 16, E109. https://doi.org/10.5888/pcd16.190109

24. Zee, B., Lee, J., Lai, M., Chee, P., Rafferty, J., Thomas, R., & Owens, D. (2022). Digital solution for detection of undiagnosed diabetes using machine learning-based retinal image analysis. *BMJ Open Diabetes Research & Care*, 10(6), e002914. https://doi.org/10.1136/bmjdrc-2022-002914