LLM-Powered Prediction of Hyperglycemia and Discovery of Behavioral Treatment Pathways from Wearables and Diet

Abdullah Mamun, Asiful Arefeen, Susan B. Racette, Dorothy D. Sears, Corrie M. Whisner, Matthew P. Buman, Hassan Ghasemzadeh*

Presented by: Abdullah Mamun

Date: July 9, 2025

abdullah-mamun.com

arxiv.org/pdf/2503.03935

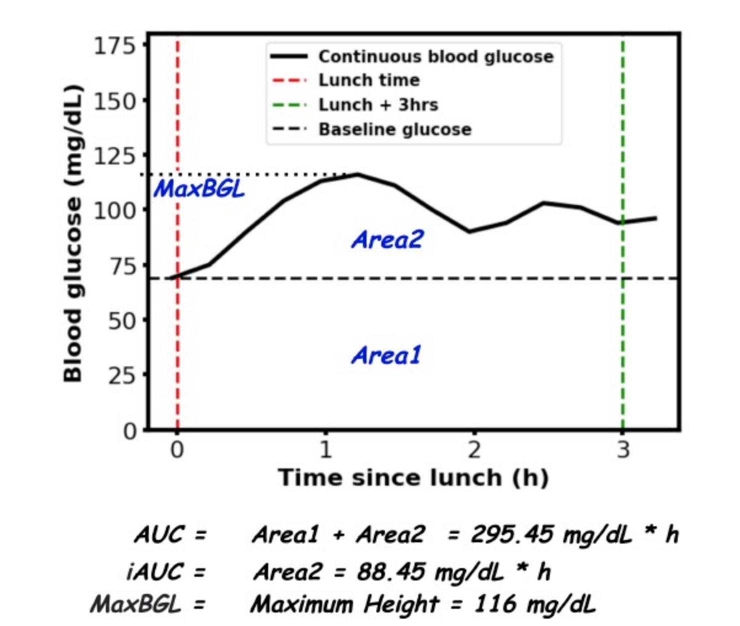
Introduction to Postprandial Hyperglycemia and Prediction

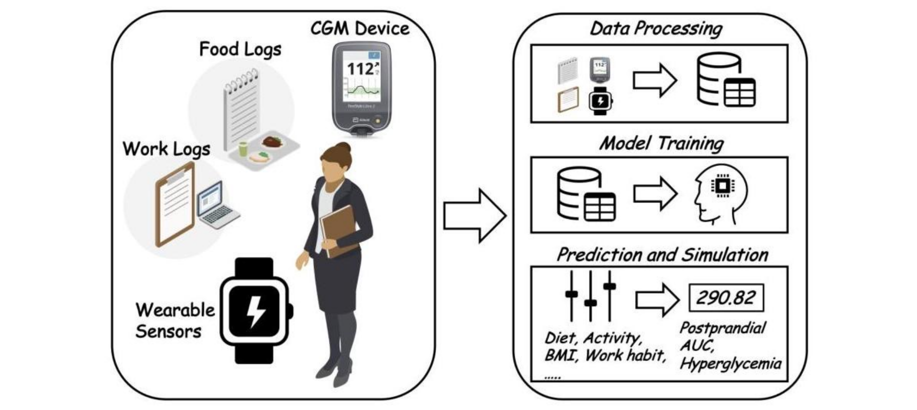
Significance of Postprandial Hyperglycemia

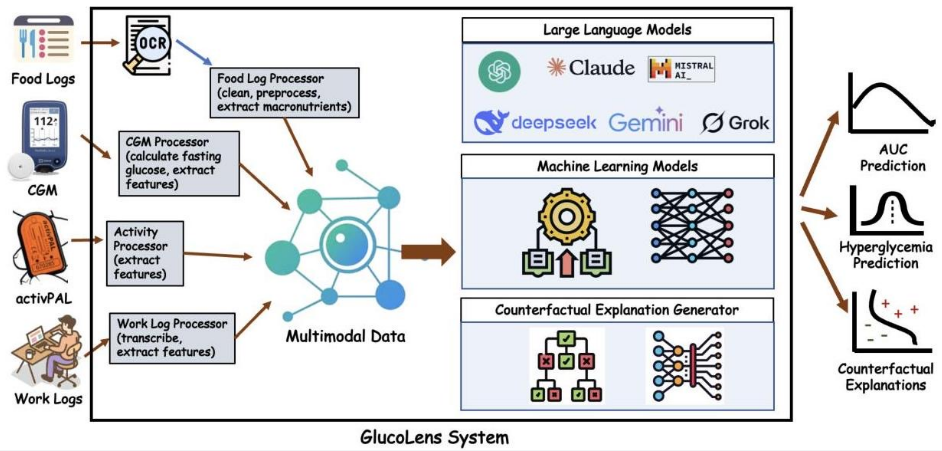
Postprandial hyperglycemia, defined by elevated blood glucose after meals, is a critical marker for progression toward type 2 diabetes. The postprandial area under the curve (AUC) is an important metric for blood glucose regulation and potential diabetes risk assessment.

Opportunity for Prediction and Intervention

By forecasting postprandial AUC in advance using lifestyle information such as diet and physical activity, individuals can proactively adjust their behaviors to maintain healthy glucose levels, potentially preventing the onset of diabetes.



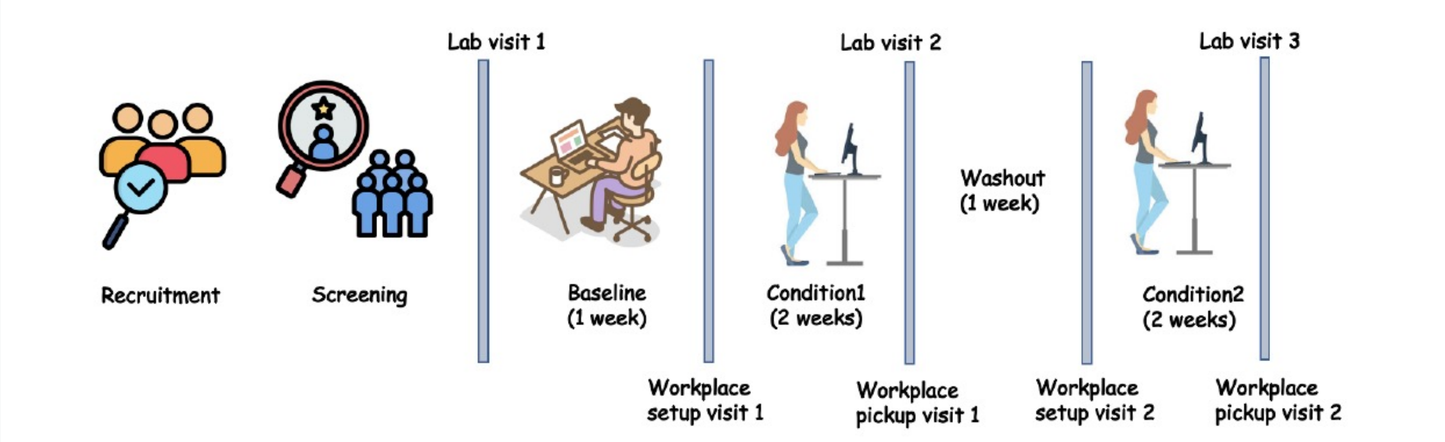




GlucoLens: An Explainable ML Solution

- GlucoLens is an explainable machine learning system designed to predict postprandial AUC and hyperglycemia.
- Integrates advanced data processing, LLMs, and trainable ML models.
- Inputs include continuous glucose monitoring (CGM), physical activity tracked by wearable devices, and detailed food and work logs.

Clinical Trial and Data Collection



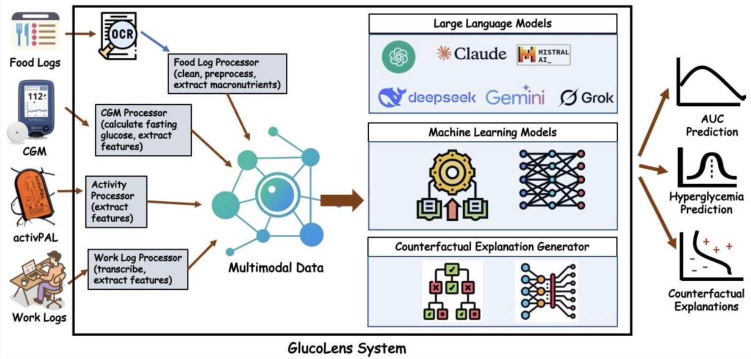
WorkWell Study Overview

- A five-week clinical trial, involved 10 full-time working adults.
- Data from CGM devices, activPAL, GENEActiv, food logs, and work logs
- Lunches were standardized and their nutritional contents precisely tracked.

Lifestyle and Activity Interventions

Participants underwent Baseline (usual habits), 'Stand' (maximal standing), and 'Move' (maximal movement) conditions in randomized order

(a) activPAL micro device, to be worn on thigh (b) Freestyle Libre CGM patch and reader (c) GENEActiv wearable wristband



No.	*Feature name/ shorthand*	*Sensor +GL*	*Sensor +Macro*	*Self +GL*	*Self +Macro*	*All*
1	Fasting glucose	☒	☒	☒	☒	☒
2	Recent CGM	☒	☒	☒	☒	☒
3	Lunch time	☒	☒	☒	☒	☒
4	Work from home	☒	☒	☒	☒	☒
5	BMI	☒	☒	☒	☒	☒
6	Calories	☒	☒	☒	☒	☒
7	Calories from fat	☒	☒	☒	☒	☒
8	Saturated fat	☒	☒	☒	☒	☒
9	Trans fat	☒	☒	☒	☒	☒
10	Cholesterol	☒	☒	☒	☒	☒
11	Sodium	☒	☒	☒	☒	☒
12	Total carbs	☒	☒	☒	☒	☒
13	Sugar	☒	☒	☒	☒	☒
14	Work start time	☒	☒	☒	☒	☒
15	Day of the week	☒	☒	☒	☒	☒
16	activPAL	☒	☒	☐	☐	☒
17	Self reported acitivity	☐	☐	☒	☒	☒
18	GL	☒	☐	☒	☐	☒
19	Net carbs	☐	☒	☐	☒	☒
20	Fat	☐	☒	☐	☒	☒
21	Protein	☐	☒	☐	☒	☒
22	Fiber	☐	☒	☐	☒	☒

Feature Engineering and Data Processing

1 Multimodal Data Processing

- Handwritten food and work logs were digitized using OCR and manual intervention.
- Features were engineered from dietary macronutrients, glycemic load calculations, wearable-derived activity metrics, and work habits.
- CGM data were processed for fasting and recent glucose metrics.

2 Comprehensive Feature Sets

Five different feature sets were formed, integrating self-reported and sensor-based activity data, macronutrients, glycemic load, and daily routines, yielding 31 features for model input including both objective and self-reported lifestyle information.

Machine Learning Architectures and Modeling Approaches

Backbone Model Experiments

GlucoLens utilized multiple ML model backbones, including Random Forests (RF), Ridge Regression, Multilayer Perceptrons (MLP), XGBoost, and TabNet. Extensive hyperparameter tuning was performed, including variation in model depth, regularization, and ensemble combinations.

Integration with Large Language Models (LLMs)

Zero-shot LLMs (e.g., GPT-4, Claude Opus 4) were employed for predictions and as hybrid inputs to ML models. LLM-only, hybrid, and base models were systematically compared for performance and interpretability.

Target outcomes	AUC, MaxBGL, Hyperglycemia
Feature sets	Sensor + Macro, Self + Macro, Sensor + GL, Self + GL, All
Predictors	RF, Ridge, MLP, XGBoost, TabNet, GPT-3.5, GPT-4, Mistral Large, Gemini Flash 2.0, Claude Opus 4, Grok 3, Deepseek V3, Gly_Hybrid, Gly_Hybrid_v2, Gly_Max, Hybrid Predictors for Classification (RF+MLP, RF+XGB, XGB+MLP, RF+XGB+MLP).
Ridge variations	$\alpha \in \{1, 0.1, 0.01\}$
RF variations	$n_{est} \in \{10, 50, 100\}$
MLP variations	13 variations; see Table 3

Prompt

Instruction:
The goal is to predict the 3-hour postprandial AUC (area under the CGM curve from lunch to 3 hours after lunch, not the incremental AUC) based on the following features:

['fasting_glucose', 'recent_cgm', 'lunch_time', 'work_at_home', 'recent_activity', 'bmi', 'Calories', 'Calories From Fat', 'Total Fat (g)', 'Saturated Fat (g)', 'Trans Fat (g)', 'Cholesterol (mg)', 'Sodium (mg)', 'Total Carbs (g)', 'Fiber (g)', 'Sugars (g)', 'Net Carbs(g)', 'Protein (g)', 'is_Friday', 'is_Monday', 'is_Thursday', 'is_Tuesday', 'is_Wednesday', 'sitting_total', 'standing_total', 'stepping_total', 'sitting_at_work', 'standing_at_work', 'stepping_at_work', 'work_start_time', 'glycemic_load'].

fasting_glucose and recent_cgm are given in mg/dL. lunch_time and work_start_time are represented as hour values (e.g., 7.75 means 7:45 AM, 13.50 means 1:30 PM). recent_activity score is calculated by taking the average percentage of time spent in walking activity in the previous days of the same phase and adding 0.5 times the average percentage of time spent in standing activity in the previous days of the same phase. sitting, standing, and stepping features are in seconds for the specific day before lunch.

Task:
Predict the 3-hour postprandial AUC for the given features.
Give me just the number enclosed within the <Prediction></Prediction> tags.

Input:

[48.0, 58.0625, 12.25, 1.0, 10.0, 36.7, 350.0, 100.0, 12.0, 2.0, 0.0, 45.0, 220.0, 27.3, 5.0, 3.0, 22.3, 32.3, 0.0, 0.0, 0.0, 0.0, 1.0, 17363.8, 1393.7, 380.1, 16843.6, 132.2, 124.2, 7.5, 14.7641].

Output:

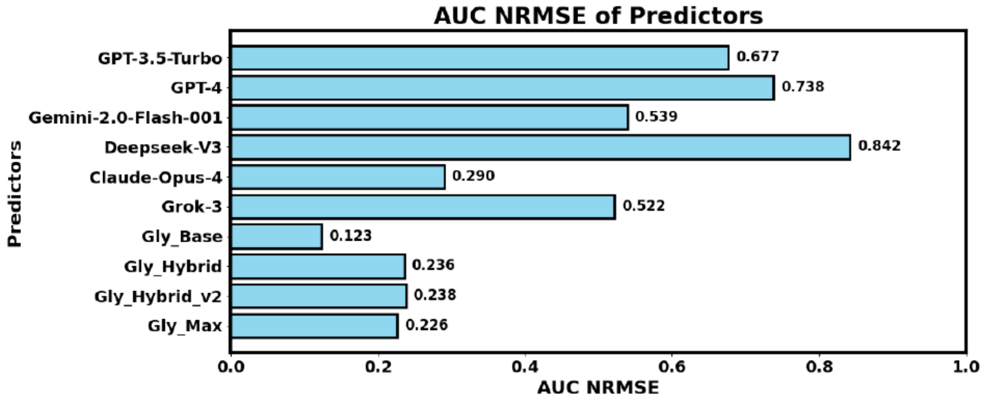
Results: AUC Prediction Performance

Table 4. Normalized Root Mean Squared Errors (NRMSE) of our GlucoLens models (RF, Ridge, MLP, XGBoost, TabNet) for different feature sets in the prediction of postprandial AUC. Explanations of the feature sets can be found in Table 1.

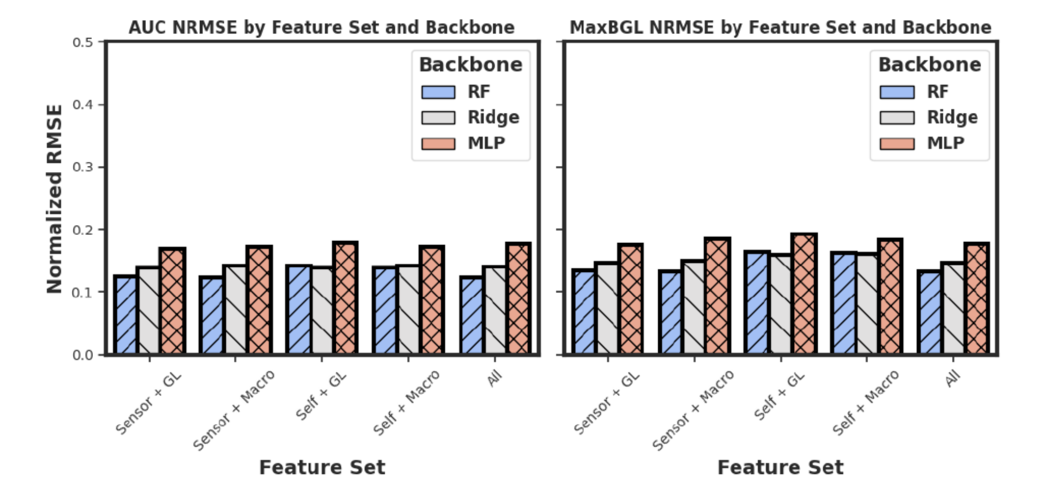
Feature Set	*RF*	*Ridge*	*MLP*	*XGBoost*	*TabNet*
Sensor + GL	**0.125**	0.139	0.169	0.137	0.160
Sensor + Macro	**0.123**	0.142	0.172	0.139	0.147
Self + GL	0.142	**0.139**	0.178	0.152	0.154
Self + Macro	**0.139**	0.142	0.172	0.149	0.151
All	**0.123**	0.140	0.176	0.137	0.151

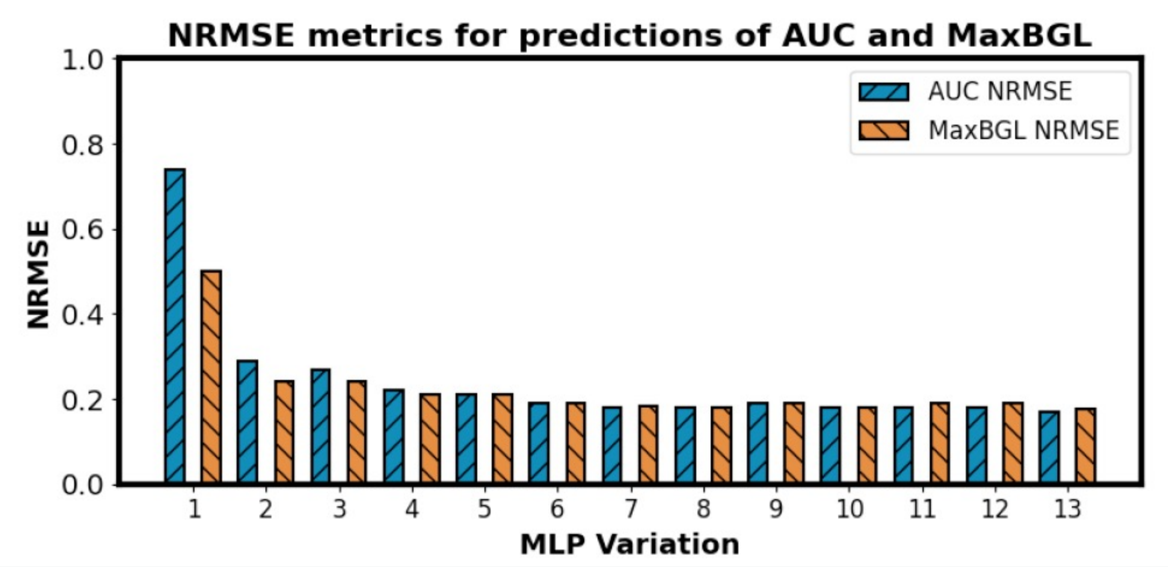
Table 5. AUC NRMSE results of different variations of our solution. Gly_Base = GlucoLens regressor with no LLM, Gly_LLM = LLM only prediction (zero-shot) after multimodal data processing by GlucoLens. The hybrid predictors are explained in Table 2.

Backbone	*Gly_Base*	*Gly_LLM*	*Gly_Hybrid*	*Gly_Hybrid_v2*	*Gly_Max*
RF	**0.123**	0.290	0.241	0.238	0.226
XGBoost	0.137		0.236	0.242	0.259

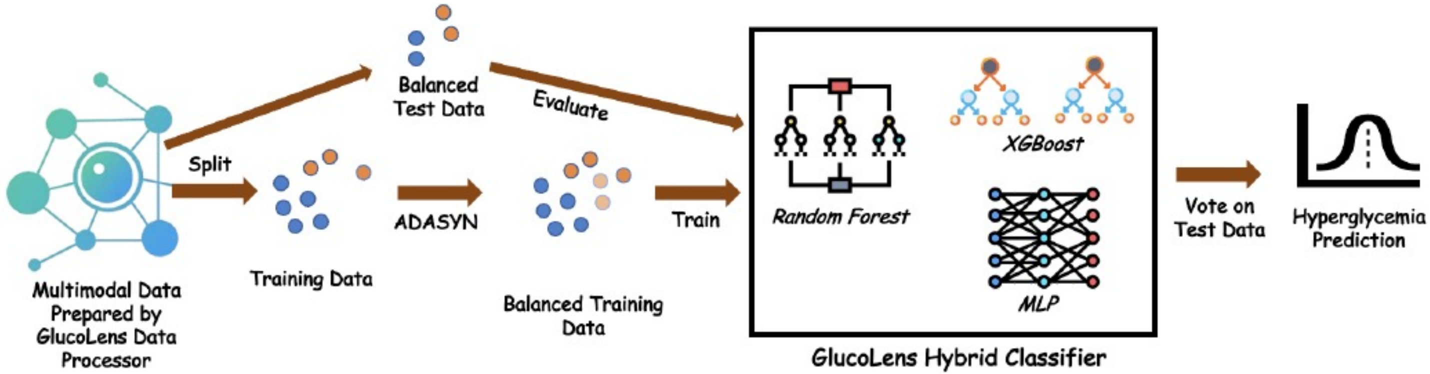


Results: MaxBGL and MLP performances





Hyperglycemia Detection Models



Hyperglycemia Detection Results

Classifier	*Accuracy*	*Precision*	*Recall*	*F1*
RF	0.698	0.737	0.699	0.685
XGB	0.685	0.720	0.692	0.682
MLP	0.620	0.626	0.620	0.589
RF+XGB	0.695	0.730	0.695	0.683
RF+MLP	0.668	0.700	0.668	0.650
XGB+MLP	0.687	0.712	0.687	0.672
RF+XGB+MLP	**0.712**	**0.740**	**0.712**	**0.702**

Size of training set	*Accuracy*	*Precision*	*Recall*	*F1*
70% training, 30% test	0.674	0.706	0.674	0.660
80% training, 20% test	0.660	0.729	0.702	0.690
87% training, 13% test	0.712	0.740	0.712	0.702
90% training, 10% test	0.717	0.744	0.717	0.705
95% training, 5% test	**0.733**	**0.751**	**0.733**	**0.716**
99% training, 1% test	0.730	0.625	0.730	0.660

Hyperglycemia Detection Results

Orignal exmple: Hyperglycemia

Current feature values:
Fiber: 1 g
Stepping duration: 8.95 minutes

Counterfactual examples: Normal blood glucose level

Option 1: Increase fiber intake to 5 grams ↑.

Option 2: Increase stepping duration to 39.38 minutes ↑.

Orignal exmple: Normal blood glucose level

Current features values:
Work start time: 11 AM
Sitting at work: 48.31 minutes
Lunch time: 1 PM
Calories in lunch: 780 kCal

Counterfactual examples: Hyperglycemia

Option 1: Start working at 6 AM ↓, increase sitting duration at work to 148.62 minutes ↑, eat lunch at 12 PM ↓, increase lunch calories to 827 kCal ↑.

Thank You!
Questions?

abdullah-mamun.com

a.mamun@asu.edu

𝕏 : @AB9Mamun

arxiv.org/pdf/2503.03935