Forewarning Postprandial Hyperglycemia with Interpretations using Machine Learning

Asiful Arefeen, Samantha Fessler, Carol Johnston, and Hassan Ghasemzadeh

***Abstract*—Postprandial hyperglycemia (PPHG) is detrimental to health and increases risk of cardiovascular diseases, reduced eyesight, and life-threatening conditions like cancer. Detecting PPHG events before they occur can potentially help with providing early interventions. Prior research suggests that PPHG events can be predicted based on information about diet. However, such computational approaches (1) are data hungry requiring significant amounts of data for algorithm training; and (2) work as a black-box and lack interpretability, thus limiting the adoption of these technologies for use in clinical interventions. Motivated by these shortcomings, we propose, *DietNudge*[1], a machine learning based framework that integrates multi-modal data about diet, insulin, and blood glucose to predict PPHG events before they occur. Using data from patients with diabetes, we demonstrate that our model can predict PPHG events with up to 90% classification accuracy and an average F1 score of 0.93. The proposed decision-tree-based approach also identifies modifiable factors that contribute to an impending PPHG event while providing personalized thresholds to prevent such events. Our results suggest that we can develop simple, yet effective, computational algorithms that can be used as preventative mechanisms for diabetes and obesity management.**

***Index Terms*—Machine learning, decision tree, diabetes, continuous glucose monitor, postprandial hyperglycemia.**

I. Introduction

Postprandial hyperglycemia (PPHG) is characterized by hyperglycemic spikes in blood glucose level. As defined by the American Diabetes Association (ADA) and the World Health Organization (WHO), the threshold for unacceptable postprandial glycemia is 8.89 mmol/L (>160 mg/dL) at any time after the meal [1]. Long-term exposure to hyperglycemia reduces glycemic control and enhances the development of cancer, macrovascular complications, cerebrovascular and cardiovascular diseases [2], [3]. The consequences are more severe among individuals living with obesity and diabetes. Therefore, the significance of predicting PPHG events, even before consuming a meal, is evident.

Continuous glucose monitors (CGM) are utilized for effective diabetes management as they transmit blood glucose concentration at a consistent frequency over extended time spans. Although CGM sensors have shown efficacy in reliably estimating blood glucose level in real-time, they are not equipped with computational algorithms to predict and warn the users of PPHG. Hence, developing an algorithm that predicts and conveys information to the users regarding imminent PPHG events and underlying modifiable factors remains unexplored. We hypothesize that blood glucose data transmitted by advanced wearables can be used to develop a machine learning model to accurately predict hyperglycemia.

Several prior studies attempted to develop algorithm for predicting glycemic response [4]–[6]. In particular, Karim et al. [7] forecast the glycemic response following the consumption of a meal by providing a feed-forward network with information including applied bolus insulin dose, baseline blood glucose concentration, maximum carbohydrate absorption rate, area under the carbohydrate absorption curve, and time elapsed since the last basal insulin. Prendin et al. [8] developed and compared thirty linear and non-linear predictive models by forecasting glucose concentrations and hypoglycemic events. Gu et al. [9] leveraged food records, drug and insulin intake, physical activities and sleep quality captured by smartphone sensors and a deep RNN (Recurrent Neural Network) model to predict abnormal glucose events such as hyperglycemia and hypoglycemia. Although the aforementioned methods can detect PPHG events, they fail to provide interpretable results or specific preventive measures and guidelines to prevent hyperglycemia. Additionally, prior studies developed data-hungry and computational expensive models, making such more less appealing for implementation on embedded systems and mobile devices.

Therefore, we observe clear gaps in the literature and to bridge them, we propose, *DietNudge*, a computationally-simple and human-readable machine learning model to predict hyperglycemia and counsel users with necessary modifiable factors and decision support for controlling blood glucose. We make all the resources and software code of DietNudge public for future aspirants to reproduce the work.

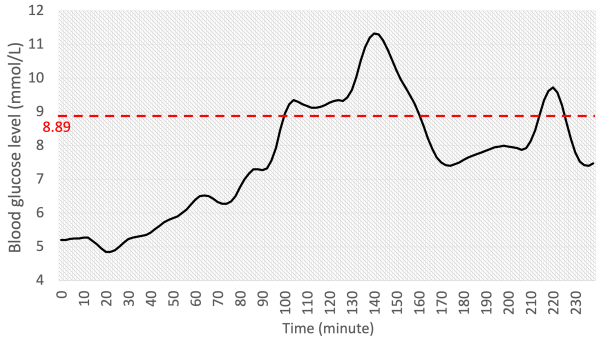


Fig. 1: Glucose concentration captured using a CGM. Postprandial glycemic response exceeds threshold for hyperglycemia.

This work was supported in part by the National Science Foundation, under grants CNS-2210133, CNS-2227002, IIS-1954372, and IIS-1852163.

Authors are with the College of Health Solutions, Arizona State University, Phoenix, AZ 85004, USA email: aarefeen@asu.edu

[1]Resources available at: https://github.com/Arefeen06088/DietNudge

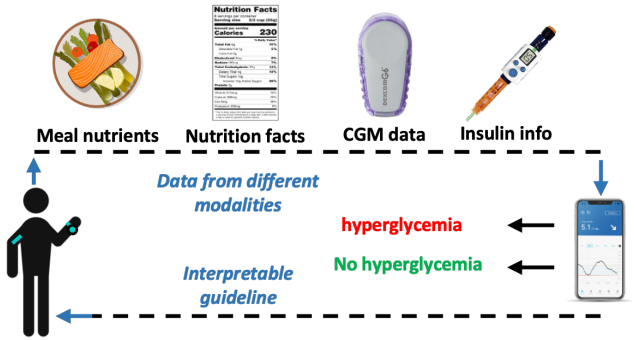


Fig. 2: DietNudge accepts data about various sources such as a camera, wearables, and user input to forecast PPHG.

II. System Design

As shown in Fig. 2, DietNudge integrates data about diet, insulin, and baseline glucose level from various sources. The machine learning algorithm takes these data as input and forecasts the incidence of an impending PPHG event. The data are fed into the algorithm by user or passive sensors; e.g., data derived from user scans of food items or nutrition fact labels, blood glucose data transmitted by CGM attached to the body and data from a Bluetooth enabled insulin pen or insulin pump to forecast hyperglycemia upon what the user is about to eat.

Since personal health data, including CGM data, are sensitive, they are not readily available in plenty. Therefore, the notion is that decision tree models would be a good fit for analyzing personal glucose data as they perform well with smaller datasets [10]. Furthermore, unlike SVM and neural networks, which follow non-linear decision boundaries, decision trees explore linear decision boundaries for classification and regression tasks and produce threshold values, which are easily interpretable for behavioral modification.

Assuming we have n instances each with D features from a meal (e.g., carb, fat, fiber) and health conditions (e.g., insulin amount, baseline glucose level), the i^{th} sample would be, $\mathbf{x}_i = [x_i^{(1)}, x_i^{(2)}, ..., x_i^{(D)}]$ yielding a D-dimensional feature set $\mathbf{x} \in \mathbb{R}^{n \times D}$. Our goal is to predict, given his/her health conditions, if a user will encounter hyperglycemia upon consuming the meal. Hence, our target domain of i^{th} sample $\mathbf{y}_i$ is a binary variable to predict and yields a target vector $\mathbf{y} \in \{0,1\}^{n \times 1}$.

Decision trees require decision making, i.e., whether to break a leaf node into more branches based on their impurity. The two most common impurity measures used in decision tree algorithms are Entropy and Gini Index. Given probability P_j for j^{th} class, Entropy is given by

$$E = -\sum_{j=1}^{n} P_j \cdot log(P_j)$$

and Gini Index is calculated as

$$G = 1 - \sum_{j=1}^{n} {P_j}^2$$

Because we aim to develop both personalized and population-based models, we used both Entropy and the Gini Index for impurity measurement. Other hyper parameters like maximum tree depth and minimum sample counts for splitting a node are chosen on a trial-and-error basis to maximize the prediction accuracy.

When training a machine learning algorithm, we use mean accuracy and μ-average F1 score as our evaluation metrics because performance of the model is iteratively measured across varying number of samples under both personalized and population-wise setups. The μ-average F1 score can be expressed as the harmonic mean of μ-average precision and μ-average recall and is given by

$$\overline{F_1}_{\mu} = \frac{2\ \overline{p}_{\mu} \overline{r}_{\mu}}{\overline{p}_{\mu} + \overline{r}_{\mu}}$$

where, μ-average precision is written as

$$\overline{p}_{\mu} = \frac{TP_1 + TP_2 + ...}{TP_1 + TP_2 + ... + FP_1 + FP_2 + ...}$$

and, μ-average recall is given by

$$\overline{r}_{\mu} = \frac{TP_1 + TP_2 + ...}{TP_1 + TP_2 + ... + FN_1 + FN_2 + ...}$$

The true positive values (TP_1,TP_2,...), false positive numbers (FP_1,FP_2,...), and false negative numbers (FN_1,FN_2,...) are derived from the model's prediction on test samples (i.e., test meals and test subjects) in personalized and population-based setups, respectively.

We hypothesize that factors leading to hyperglycemia are subject dependent. A subset of factors could be highly correlated to one subject's PPHG and could be loosely correlated to that of others. Therefore, we employ Recursive Feature Elimination (RFE) to identify the top contributors in hyperglycemia forecasting. In general, RFE forms a classification/regression model and initiates a k-fold cross-validation to recursively evaluate smaller and smaller feature sets. On each iteration, RFE identifies the least important feature and ranks the features accordingly. The iteration continues until a desired number of top contributing features are identified. In our case, although we obtain factors leading to PPHG using the decision tree enroute to PPHG events, we also use a decision tree based RFE to identify the top contributors on a personal level.

III. Experimental Validation

A. Dataset

We used the Nutrients Absorption dataset [11], which include data from five participants (4 T2DM and 1 T1DM patients, 4 male and 1 female, mean age 51.6 years), who were fed a combined total of 167 meals. In addition to meal nutrient amounts (carb, fat, and fiber), the dataset contains injected bolus insulin amounts, time elapsed since last insulin dose (DFB), blood glucose concentration before each meal (SBGL), and four hours of CGM data following the consumption of meals with S7 EasySense CGM System by Medtrum.

B. Model Construction

For PPHG forecasting, two different model types were developed using the decision tree algorithm: leave-one-meal-out (LOMO), which is subject-dependent and leave-one-subject-out (LOSO), which is subject-independent. The subject-dependent (LOMO) model was obtained using data from all the meals except one, which was left out for testing, and

TABLE I: Comparison of Decision Tree (DT) and SVM models with subject-dependent and subject-independent frameworks when fiber is included. On average, we received **76.2** and **79.7%** accuracy respectively across 5 subjects with LOMO setup.

	DT - LOMO					SVM - LOMO					DT - LOSO	SVM - LOSO
Metric type	**P1**	**P2**	**P3**	**P4**	**P5**	**P1**	**P2**	**P3**	**P4**	**P5**		
Mean accuracy (%)	89.6	74.4	61.5	70	85.7	72.4	100	54.8	100	71.4	68.8	67.5
μ-average F1 score	0.93	0.81	0.67	0.78	0.9	0.84	1	0.7	1	0.83	0.81	0.81

TABLE II: Subject-dependent (LOMO) and subject-independent (LOSO) performance of DT and SVM models when fiber is excluded. Note that, average prediction performance slightly improves and extends upto **79.6%** accuracy across 5 subjects for DT-LOMO when fiber is excluded. Depending on the dataset, the prediction horizon of our approach is 4 hours.

	DT - LOMO					SVM - LOMO					DT - LOSO	SVM - LOSO
Metric type	**P1**	**P2**	**P3**	**P4**	**P5**	**P1**	**P2**	**P3**	**P4**	**P5**		
Mean accuracy (%)	89.9	80	65.3	77.1	85.7	72.5	74.4	54.8	62	72.1	70	70
μ-average F1 score	0.93	0.83	0.71	0.8	0.9	0.84	0.85	0.7	0.76	0.83	0.81	0.81

this was repeated across all the meals and subjects. As such, all subjects had a personalized PPHG forecasting model. In contrast, the subject-independent (LOSO) model was obtained using data and features from all subjects except one, which was set aside for testing. This was repeated across all subjects, leading to one generalized model for all participants. This way, we evaluated if subject-dependent or subject-independent models give us better results in PPHG forecasting.

C. Performance

We split our results into two branches based on inclusion and exclusion of consumed fiber amount. We also present the performance of a SVM-based (Support Vector Machine) PPHG forecasting model for a valid comparison.

Table I presents results including fiber amounts while Table II shows results when fiber amounts were excluded from the dataset. As expected the personalized models performed better than subject-independent models. For example, decision tree and SVM based subject-independent models offered the best accuracy at 70% with a μ-average F1 score of 0.81. In contrast, personalized SVM models for P2 and P4 recorded maximum accuracy and μ-average F1 score.

Although subject-dependent SVM models achieved the highest accuracy for P2 and P4, subject-dependent decision tree models achieved a better performance overall. Accuracy values and F1 scores were lowest for P3 be it a decision tree model or SVM. However, it is worth noting that P3 was the oldest of five participants and had the least number of meal records or data points, which are likely responsible for the abnormality in forecasting results. Forecasting horizon was 4 hours as imposed by the dataset.

D. Top factors

We further employed RFE algorithm on the data to get top two influential factors in PPHG forecasting. For subject-specific models, carb amounts and carb composition were frequently identified as strongest predictors of PPHG. Table III shows other variables including insulin amounts, fat amounts and baseline blood glucose also played leading roles for forecasting some participants' hyperglycemia. Carb composition and DFB were dominant in generalized models. This finding is consistent with prior research [12] which suggests that altering the type and/or amount of dietary carb can improve postprandial glucose.

TABLE III: Top 2 driving factors of PPHG.

Validation method	**Participant no.**	**Top 2 contributing features**
LOMO	P1	Insulin, Carbs amount
	P2	Carbs amount, Fat amount
	P3	Carbs amount, Carbs composition
	P4	Insulin, Carbs amount
	P5	SBGL, Carbs composition
LOSO	Generalized model	Carbs composition, DFB

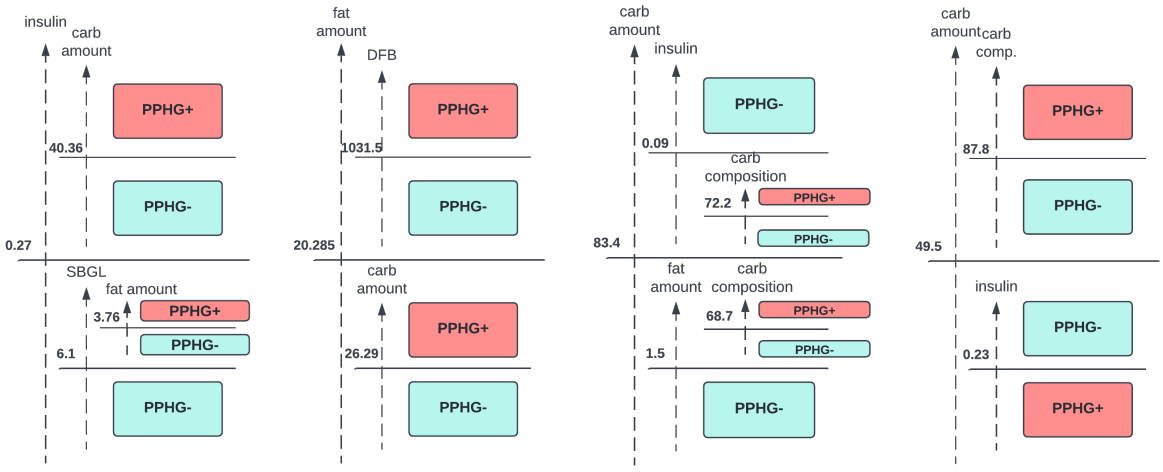
E. Interpretation

Since decision trees follow linear decision boundaries and generate understandable thresholds for each variable, they can be utilized for hyperglycemia, diabetes or obesity management. DietNudge presents those factors and corresponding thresholds in a way that they become useful, informative and may operate as guidelines for users.

Fig. 3 shows factors with their thresholds for subjects, P1, P2, P3 and P4. P3, for example, might be able to avoid hyperglycemia if the consumed meal contains carb less than 83.4 grams and fat less than 1.5 grams or if overall carb composition is below 68.7%. From a different standpoint, if a consumed meal contains carb amount greater than 83.4 grams, a basal insulin dose higher than 0.09 might help evade a PPHG event.

IV. Conclusions and Future Work

We developed a simple interpretable machine learning model for forecasting hyperglycemia in patients with diabetes. The developed decision tree model provides insights about effective interventions to prevent PPHG events. Our results demonstrate the feasibility of designing computational algorithms trained with small amounts of training data for



(a) Interpretation for P1 (b) Interpretation for P2 (c) Interpretation for P3 (d) Interpretation for P4

Fig. 3: Interpretation of decision tree results for hyperglycemia management using threshold values produced by DietNudge. Under LOMO setup, decision trees are generated as many times as the available number of meals, however, most are identical. Most recurring ones are taken to the next step for behavioral modification.

accurate prediction of important health events. We showed that carb component of the diet (i.e., carb amount and carb composition) is a strong predictor of PPHG event, however, the thresholds that trigger a health event are not universal across all patients. This finding emphasizes the importance of personalized interventions.

We recognize that hyperglycemia depends on many other factors such as physical activity choices, poor disease management, non-diabetes medications, or skipping glucose-lowering medication [13]. To the best of our knowledge, however, current publicly available datasets do not contain such a comprehensive set of variables. For this research, our analysis was limited to designing machine learning models based on baseline glucose level, diet, and insulin amount only. Our ongoing and future work involves construction of comprehensive dataset that integrates different modalities from a large cohort of individuals as well as the development of interpretable machine learning models for PPHG forecasting and blood glucose management.

V. Acknowledgement

This work was supported in part by the National Science Foundation, under grants CNS-2210133, CNS-2227002, IIS-1954372, and IIS-1852163. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding organizations.

References

[1] J. C. Brand-Miller, K. Stockmann, F. S. Atkinson, P. Petocz, and G. Denyer, "Glycemic index, postprandial glycemia, and the shape of the curve in healthy subjects: analysis of a database of more than 1,000 foods." *The American journal of clinical nutrition*, vol. 89 1, pp. 97–105, 2009.

[2] D. M. Nathan, J. B. Buse, E. S. Horton, H. E. Lebovitz, R. Rizza, M. M. Engelgau, T. P. Erlinger, D. E. Goldstein, and F. R. Kaufman, "Postprandial blood glucose. american diabetes association." *Diabetes care*, vol. 24 4, pp. 775–8, 2001.

[3] A. V. dos Santos, H. F. Loponte, N. R. Mantuano, I. A. Oliveira, I. F. D. Paula, L. K. Teixeira, J. C. M. de Freitas-Junior, K. C. Gondim, N. Heise, R. Mohana-Borges, J. A. Morgado-Díaz, W. B. Dias, and A. R. Todeschini, "Hyperglycemia exacerbates colon cancer malignancy through hexosamine biosynthetic pathway," *Oncogenesis*, vol. 6, 2017.

[4] D. J. Albers, M. E. Levine, B. J. Gluckman, H. N. Ginsberg, G. Hripcsak, and L. Mamykina, "Personalized glucose forecasting for type 2 diabetes using data assimilation," *PLoS Computational Biology*, vol. 13, 2017.

[5] T. Hamdi, J. B. Ali, V. D. Costanzo, F. Fnaiech, E. Moreau, and J.-M. Ginoux, "Accurate prediction of continuous blood glucose based on support vector regression and differential evolution algorithm," *Biocybernetics and Biomedical Engineering*, vol. 38, pp. 362–372, 2018.

[6] D. A. Zeevi, T. Korem, N. Zmora, D. Israeli, D. Rothschild, A. Weinberger, O. Ben-Yacov, D. Lador, T. Avnit-Sagi, M. Lotan-Pompan, J. Suez, J. A. Mahdi, E. Matot, G. Malka, N. Kosower, M. Rein, G. Zilberman-Schapira, L. Dohnalová, M. Pevsner-Fischer, R. Bikovsky, Z. Halpern, E. Elinav, and E. Segal, "Personalized nutrition by prediction of glycemic responses," *Cell*, vol. 163, pp. 1079–1094, 2015.

[7] R. A. H. Karim, I. Vassányi, and I. Kósa, "Improved methods for mid-term blood glucose level prediction using dietary and insulin logs," *Medicina*, vol. 57, 2021.

[8] F. Prendin, S. D. Favero, M. Vettoretti, G. Sparacino, and A. Facchinetti, "Forecasting of glucose levels and hypoglycemic events: Head-to-head comparison of linear and nonlinear data-driven algorithms based on continuous glucose monitoring data only," *Sensors (Basel, Switzerland)*, vol. 21, 2021.

[9] W. Gu, Y. Zhou, Z. Zhou, X. Liu, H. Zou, P. Zhang, C. J. Spanos, and L. Zhang, "Sugarmate: Non-intrusive blood glucose monitoring with smartphones," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 1, pp. 54:1–54:27, 2017.

[10] D. Saravanakumar, N. Ananthi, and M. S. Devi, "An approach to automation selection of decision tree based on training data set," *International Journal of Computer Applications*, vol. 64, pp. 1–4, 2013.

[11] R. A. H. Karim, I. Vassányi, and I. Kósa, "After-meal blood glucose level prediction using an absorption model for neural network training," *Computers in biology and medicine*, vol. 125, p. 103956, 2020.

[12] J. H. O'Keefe and D. S. Bell, "Postprandial hyperglycemia/hyperlipidemia (postprandial dysmetabolism) is a cardiovascular risk factor," *The American journal of cardiology*, vol. 100, no. 5, pp. 899–904, 2007.

[13] W. D. Smith, A. G. Winterstein, T. E. Johns, E. I. Rosenberg, and B. C. Sauer, "Causes of hyperglycemia and hypoglycemia in adult inpatients." *American journal of health-system pharmacy : AJHP : official journal of the American Society of Health-System Pharmacists*, vol. 62 7, pp. 714–9, 2005.