Article

Classification of Postprandial Glycemic Status with Application to Insulin Dosing in Type 1 Diabetes—An In Silico Proof-of-Concept

Giacomo Cappon [1], Andrea Facchinetti [1], Giovanni Sparacino [1], Pantelis Georgiou [2] and Pau Herrero [2,*]

[1] Department of Information Engineering, University of Padova, 35131 Padova (PD), Italy
[2] Department of Electrical and Electronical Engineering, Imperial College London, London W5 5SA, UK
* Correspondence: pherrero@imperial.ac.uk

check for updates

Abstract: In the daily management of type 1 diabetes (T1D), determining the correct insulin dose to be injected at meal-time is fundamental to achieve optimal glycemic control. Wearable sensors, such as continuous glucose monitoring (CGM) devices, are instrumental to achieve this purpose. In this paper, we show how CGM data, together with commonly recorded inputs (carbohydrate intake and bolus insulin), can be used to develop an algorithm that allows classifying, at meal-time, the post-prandial glycemic status (i.e., blood glucose concentration being too low, too high, or within target range). Such an outcome can then be used to improve the efficacy of insulin therapy by reducing or increasing the corresponding meal bolus dose. A state-of-the-art T1D simulation environment, including intraday variability and a behavioral model, was used to generate a rich in silico dataset corresponding to 100 subjects over a two-month scenario. Then, an extreme gradient-boosted tree (XGB) algorithm was employed to classify the post-prandial glycemic status. Finally, we demonstrate how the XGB algorithm outcome can be exploited to improve glycemic control in T1D through real-time adjustment of the meal insulin bolus. The proposed XGB algorithm obtained good accuracy at classifying post-prandial glycemic status (AUROC = 0.84 [0.78, 0.87]). Consequently, when used to adjust, in real-time, meal insulin boluses obtained with a bolus calculator, the proposed approach improves glycemic control when compared to the baseline bolus calculator. In particular, percentage time in target [70, 180] mg/dL was improved from 61.98 (±13.89) to 67.00 (±11.54; $p < 0.01$) without increasing hypoglycemia.

Keywords: continuous glucose monitoring; decision support systems; machine learning; type 1 diabetes; gradient boosted trees; postprandial glycaemia

1. Introduction

Type 1 diabetes (T1D) is a lifelong condition characterized by the destruction of the pancreatic beta cells responsible for insulin production. Consequently, people affected by T1D strongly rely on exogenous insulin in order to maintain blood glucose (BG) levels within the safe range ([70, 180] mg/dL) [1] and avoid long-term complications, e.g., neuropathy, retinopathy, and micro/macro-vascular heart diseases, due to sustained hyperglycemia (BG > 180 mg/dL) [2]. However, administration of exogenous insulin is difficult to tune and an excess of it can cause hypoglycemia (BG < 70 mg/dL), which leads to dizziness, lightheadedness, fainting, and, in the most extreme cases, coma or even death.

Insulin is exogenously injected either via multiple daily injections (MDI), consisting of fast-acting insulin performed at meal-time (bolus insulin) and a long-acting dose once (or twice) per day, or via an infusion pump (i.e., continuous subcutaneous insulin infusion). Titration of such doses is usually done

according to self-monitoring of blood glucose (SMBG), diet, and physical activity, among others [1]. In particular, the calculation of the meal-time insulin bolus is usually done by applying the standard bolus calculator formula, which is commonly available in most insulin pumps and some glucose meters. However, due to unaccounted physiological events and variability, this formula is, in many cases, suboptimal and might lead to post-prandial hyper- and hypoglycemia [3]. In the past few years, exciting perspectives in the management of T1D were opened by the introduction of new technologies, such as minimally invasive sensors for real-time continuous glucose monitoring (CGM). CGM has revolutionized BG monitoring providing a measurement every 1–5 min, mitigating the need of the SMBG and greatly increasing the information on BG fluctuations [4,5]. CGM and additional data streams obtained from wearable devices, such as activity monitors, can provide important real-time information on the current status of the diabetic individual, and appear to be suitable to develop new "smarter" tools to empower people with T1D with decision support to improve the management of their condition [6,7]. In recent years, several methodologies have been proposed with the aim of optimizing the insulin bolus dose by using the subject's information at meal-time (e.g., BG level, carbohydrate intake, previous insulin bolus, activity counts, and therapy parameters), and have been demonstrated to improve glycemic control [8,9]. Furthermore, additional improvements can be achieved if, in addition to using information at meal-time, a prediction of postprandial glycemic status, i.e., future BG being too low, too high, or within a target range; is accounted for [10,11].

In the present work, we develop a classification technique usable at meal-time to predict the future glycemic status in the postprandial period (i.e., between 2 h and 6 h after meal) exploiting the knowledge of pre-prandial CGM measurements, carbohydrate intake estimates and insulin infusion recordings. In particular, the technique uses the extreme gradient-boosted tree (XGB) algorithm [12] to classify whether the postprandial glycemic status is expected to be too high, too low, or within a target range. Finally, to show the potential usefulness of the proposed classification methodology in the daily management of T1D insulin therapy, we demonstrate that, by using a simple heuristic technique, an improvement on glycemic control can be obtained by adjusting the insulin bolus according to the predicted glycemic status.

Assessment of the method is performed by resorting to clinical trials. This is a strategy widely employed by the diabetes technology scientific community to develop and evaluate new algorithms before moving on to real subjects [13–15]. In particular, here we use synthetic data generated with a state-of-the-art mathematical model of a T1D subject decision-making (T1D-DM) [16]. This model is an evolution of a software tool that received approval by the US Food and Drug Administration (FDA) as a substitute for pre-clinical animal testing [17], which in addition to the glucose-insulin-glucagon dynamics, includes intraday variability of insulin sensitivity and a behavioral model for a T1D population.

2. Method to Classify Future Postprandial Glycemic Status

2.1. Classification Criterion

To define the glycemic status in the post-prandial time window, i.e., if after the meal the individual glycemia will be in hypoglycemia, hyperglycemia, or within a target range, the CGM-based metric described by Herrero et al. [18] was employed. It consists in computing the minimum postprandial CGM (G_{min}) level within a predefined postprandial time window as:

$$G_{min} = \min_{t \in [t_m + 2h, \, t_m + 6h]} CGM(t), \tag{1}$$

where, t_m is the time of the meal. The rational for choosing a postprandial time window from two to six hours after meal time is that we are interested in computing the minimum postprandial glucose after the glucose peak time. Note that a typical glucose peak after meals is around 70 min but can potentially be longer for low-absorption meals. We then chose to end the window at six hours post meal to make

sure that the effect of the short-acting meal insulin bolus is over [19]. Finally, the assumption that basal insulin (i.e., long-acting insulin) is correctly adjusted is made.

Then, the classification target Y consists of three classes defined as follows:

$$Y = \begin{cases} C_1, & G_{min} < \text{th}_{\text{hypo}}, \\ C_2, & \text{th}_{\text{hypo}} \leq G_{min} \leq \text{th}_{\text{hyper}}, \\ C_3, & G_{min} > \text{th}_{\text{hyper}}, \end{cases} \tag{2}$$

where th_{hypo} and th_{hyper} are the hypoglycemia and hyperglycemia thresholds, respectively. Here, these thresholds were set to $\text{th}_{\text{hypo}} = 70$ mg/dL and $\text{th}_{\text{hyper}} = 140$ mg/dL. Figure 1 shows a graphical representation of the employed classification criterion.
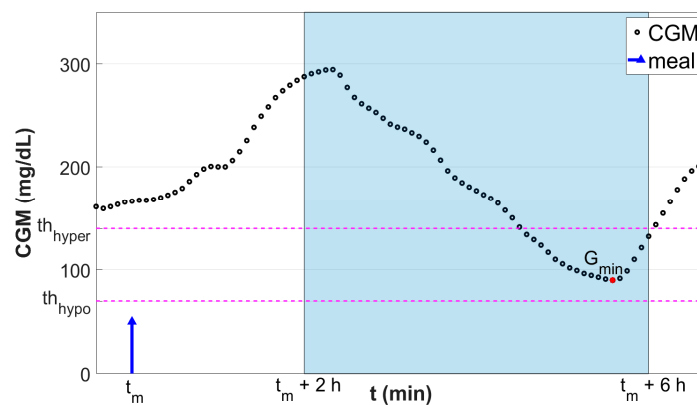

Figure 1. Graphical representation of the classification scheme. Black dots are the continuous glucose monitoring (CGM) samples. Blue stem indicates the meal event. Red dot is the minimum glucose level G_{min} reached in the fixed postprandial window ($t_m + 2$ h, $t_m + 6$ h; highlighted in light blue). Horizontal magenta lines denote the thresholds used to discretize G_{min} into the classification target Y.

In this context, an extreme gradient-boosted tree model, hereafter referred to as XGB, was chosen to classify the postprandial excursion into classes C_1, C_2, or C_3.

2.2. Extreme Gradient-Boosted Tree Model

XGB is a highly effective and efficient supervised learning algorithm that has been shown to provide state-of-the-art results in many classification tasks [12]. It is particularly well known for being robust against outliers, to work well for relatively small datasets, and to automatically handle feature selection; thus, it well fitted the purpose of this work. Moreover, XGB associates each input to the probability of belonging to each one of the three considered target classes, i.e., $p(C_1)$, $p(C_2)$, and $p(C_3)$. This feature will result in being very useful, as described later in Section 3.

For each studied subject p, we trained a different XGB model, XGB_p, the aim being accounting for inter-subject variability and personalizing the methodology at the individual level. For this purpose, a software framework (schematized in Figure 2) was built to automatically manage data preparation, model tuning, and testing.
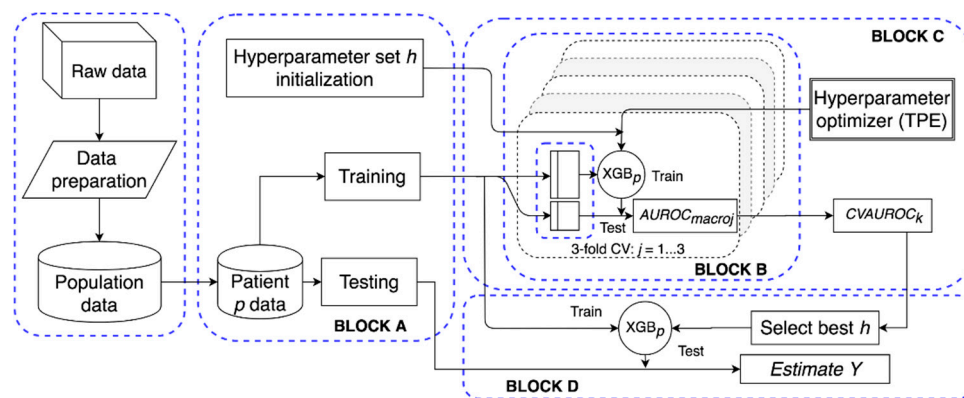
Figure 2. Structure of the proposed extreme gradient-boosted tree (XGB_p) software framework. After data preparation, for patient p, block A initializes h and splits the data in training and test; block B computes the performance of the hyperparameter set in a three-fold cross validation (CV) setting over the training set; block C implements a tree-structured Parzen estimator to optimize h; block D selects the best h set and evaluates the performance of XGB_p on the test set.

2.2.1. Features Vector and Data Preparation

The selected input (features vector; X_{ip}) for the XGB model is composed by the following entities:

- Estimated amount of ingested carbohydrates (CHO_i);
- Meal insulin bolus (IMB_i);
- Two binary indicators denoting whether there was a hypo/hyperglycemic event in the last three hours. This feature allows us to capture the physiological response to hypoglycemia (e.g., secretion of glucagon) or the ingestion of rescue carbohydrates;
- The hour-of-day of t_{mi} and three binary indicators representing the meal type (i.e., breakfast, lunch, or dinner), which are used to capture the subject's intra-day variability (e.g., circadian rhythms);
- Two features describing the time elapsed since the last insulin bolus and meal intake, respectively. This feature might help to capture specific patient behaviors, such as using multiple boluses to treat the same meal and/or snacking pattern;
- CGM data within the time window ($t_{mi} - 1$ h, t_{mi}); in addition, data were preprocessed in order to obtain additional features. In detail, for each ingested meal at time t_{mi}, CGM data, the estimated amount of carbohydrates (CHO), and insulin data (INS), were considered within the time window ($t_{mi} - 1$ h, t_{mi}). Then, such data were processed as follows:
- CGM was used to obtain the corresponding glucose rate of change, static risk (SR), and dynamic risk (DR) [20] time series, which empower the model with additional features that capture the dynamics of the CGM signal (e.g., glycemic variability);
- CHO was used to calculate the rate of glucose appearance in the blood (Ra) within (t_{mi}, t_{mi} + 1 h) through the use of a gastrointestinal model [21] to describe carbohydrate digestion and glucose absorption;
- INS data were transformed into two continuous signals representing an estimate of plasma insulin concentration (IP) [22] and the insulin-on-board (IOB) [23] to account for insulin absorption and clearance. As per the Ra signal, IP and IOB were estimated within (t_{mi}, t_{mi} + 1 h) assuming no additional insulin infusion was in that period.

2.2.2. Model Tuning and Testing

For each subject, training of XGB_p was performed through the gradient descent algorithm [24]. Figure 2 schematizes the model tuning and testing procedure. In detail, referring to Figure 2, Block A splits the dataset (X_{ip}, Y_{ip}) into training and testing data. Block A was also in charge of initializing the hyper-parameters to random values. These hyper-parameters are: The number of trees, the maximum

depth of each tree, the subsample ratio of the training instances, the L2 regularization term on weights, and the learning rate. Then, in Block B, training data were used to tune the model and, given a set of hyper-parameters h, performance was assessed in a three-fold cross validation setting. Specifically, training data were split in three folds, of which two folds were used for training, and the third one was used to validate and evaluate the model. Then, performance of the k-th hyper-parameter set was quantified in terms of intra-folds macro-average area under the receiver operator characteristic curve (AUROC) [25]:

$$CVAUROC_k = \frac{1}{3} \sum_{j=1}^{3} AUROC_{macro_j},\tag{3}$$

where $AUROC_{macro_j}$ denotes the intra-folds macro-average AUROC computed using the j-th fold as the validation set:

$$AUROC_{macro_j} = \sum_{c=1}^{3} AUROC_{C_c},\tag{4}$$

where $AUROC_{C_c}$ stands for the AUROC of the c-th class computed in a one-vs.-all fashion.

Hyper-parameters were optimized by iterating Block C. Specifically, it implemented the tree-structured Parzen estimator (TPE) technique [26], i.e., an optimization algorithm that, for each iteration, collects a new observation (the k-th set of hyper-parameters and the respective performance $CVAUROC_k$) and selects which set of hyper-parameters should be tried in the next iteration in order to improve the model performance. Finally, Block D selected the set of hyper-parameters k associated to the maximum $CVAUROC_k$ and it obtained the predicted classes on the test set.

2.3. Simulated Dataset

To evaluate the proposed XGB model, data of 100 virtual adult subjects were generated by simulation over two months using T1D-DM, which expands the model implemented in the University of Virginia/Padova T1D Simulator [17] with new modules describing glucose sensors error [27,28], intraday variability of insulin sensitivity [29] and patient behavior. Of note, physical activity was not included in the simulations, since the current version of the University of Virginia/Padova T1D simulator does not include a mathematical model describing its impact on glucose metabolism. More specifically, it consists of a large-scale maximal mathematical computer model of glucose, insulin, and glucagon dynamic of people with T1D, able to describe the inter-individual variability observed in the T1D population. It has been accepted by the FDA as a substitute to pre-clinical trials for closed-loop insulin delivery strategies and open-loop therapies [17].

The obtained dataset consists, for each individual, of one continuous time series, i.e., CGM; and two impulsive signals, i.e., INS and CHO, respectively. The first month of recordings was used to train and tune the model, while the second month was used for testing purposes. In particular, this choice allowed making a trade-off between model performance and the possibility of deploying such a model in real life. In fact, using more training data tends to improve model performance, but at the same time it requires more time to collect them, which might jeopardize patient's adherence to the therapy.

2.4. Classification Results

Figure 3 shows the AUROC distributions computed in the test set for class C_1, C_2, C_3, and their macro-average obtained in the population. On average, good performance is achieved. Specifically, the obtained average AUROC values are: $AUROC(C_1) = 0.89$ [0.86, 0.93], $AUROC(C_2) = 0.76$ [0.69, 0.82], $AUROC(C_3) = 0.86$ [0.80, 0.91], $AUROC(MACRO\text{-}AVERAGE) = 0.84$ [0.78, 0.87]. Particularly, C_2 results to be the most difficult class to predict. This is due to the discretization applied to G_{min}. Specifically, since each G_{min} is hard-assigned to a single class, the classification error is likely higher when the original G_{min} was close to the th_{hypo}/th_{hyper} thresholds before being discretized. Therefore, given that

C_2 is both upper (th_{hyper}) and lower (th_{hypo}) bounded, it is more likely that the error increases because of this discretization.
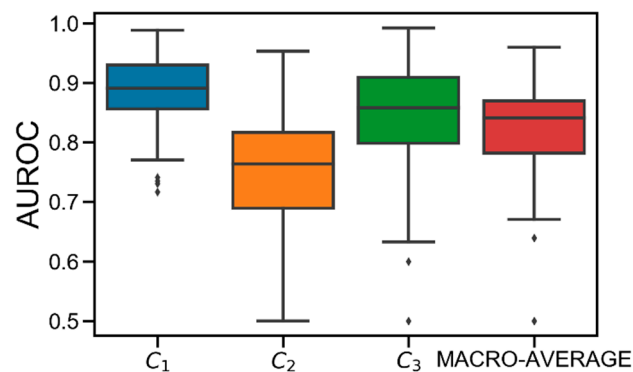
Figure 3. Boxplot representation of the distribution of area under the receiver operator characteristic curve (AUROC) obtained for C_1 (in blue), C_2 (in orange), C_3 (in green), and their macro-average (in red) obtained in the population. Black horizontal line represents median, the black box marks the interquartile range, vertical black lines are the whiskers, and black diamonds indicate outliers.

Notably, best performance is obtained for C_1, meaning that the model is accurate in detecting hypoglycemia. Finally, several critical outliers are present in C_2 and C_3. However, they are strictly higher than 0.5, meaning that XGB always behaves better than the "random" classifier [24].

Finally, Figure 4 shows the receiving operator characteristic (ROC) curves obtained for a representative subject (adult#1). The model achieves very good performance. In particular, AUROC computed for class C_1, C_2, and C_3 are 0.97, 0.83, and 0.95, respectively. Figure 4 reports the corresponding confusion matrix. Moreover, analyzing the classification error that XGB makes when it predicts the wrong class, it can be seen that it only picks adjacent classes. This is very important, since it avoids dangerous counter-actions by the patient, i.e., classifying an actual hypoglycemia as a hyperglycemia could lead the patient to increase the severity of the episode.
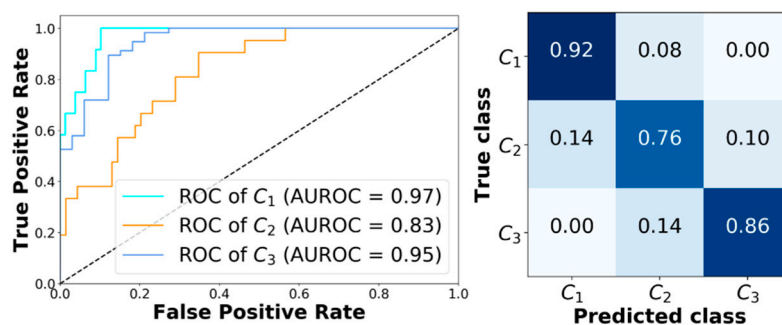
Figure 4. Classification results obtained with XGB corresponding to adult#1. Left panel: Receiving operator characteristic (ROC) curves obtained with XGB for C_1, C_2, and C_3. Right panel: The corresponding confusion matrix.

In the next section, an application of XGB is shown, the aim being to test its capability at improving glycemic outcomes, by using its outcome in real-time to adjust meal insulin boluses.

3. Application: Using the XGB Classifier to Adjust Meal Insulin Bolus

In a real-time setting, the developed XGB_p classifier can be applied at meal-time to forecast the post-prandial glycemic status (i.e., hyperglycemia, euglycemia, or hypoglycemia). Such information on the future glycemic status can be used for multiple purposes. For instance, it can be used to generate smart alerts when future adverse events are forecasted; temporarily suspend basal insulin delivery, or

suggest carbohydrate intake, to prevent hypoglycemia; and, the application which occupies us in this work, to recommend to modulate the insulin dose to be delivered at meal time.

Here, we present a proof of concept for a simple empirical strategy to adjust the meal insulin bolus according to the real-time post-prandial glucose classification provided by the XGB model at meal-time. The aim of this study was to demonstrate the potential of XGB-based strategy at improving the standard T1D insulin therapy.

3.1. Meal Insulin Dose Adjustment Strategy

At meal time t_m, if XGB classified future G_{min} belonging to C_i, IMB computed using (1) is adjusted using Equation (5), hereafter labeled as XGB-IMB:

$$IMB^*(t_m) = IMB(t_m) \times (1 + p(C_i) \times f_i(IMB(t_m))), \tag{5}$$

where $p(C_i)$ is obtained by XGB and denotes the probability of G_{min} belonging to class C_i; and $f_i(\cdot)$ is an empirical modulation function associated to C_i, which depends on the original IMB amount. In particular, $f_i(\cdot)$ is defined in Equation (6).

$$f_1(IMB(t_m)) = \begin{cases} f_{1_{LOW}} & if\ IMB(t_m)\ <\ 5\ U \\ f_{1_{LOW}}\ +\ \frac{1}{3}\Delta f_1 & if\ 5 \le IMB(t_m)\ <\ 10\ U \\ f_{1_{LOW}}\ +\ \frac{2}{3}\Delta f_1 & if\ 10 \le IMB(t_m)\ <\ 15\ U \\ f_{1_{HIGH}} & if\ IMB(t_m)\ \ge\ 15\ U \end{cases}$$

$$f_2(IMB(t_m)) = 0 \tag{6}$$

$$f_3(IMB(t_m)) = \begin{cases} f_{3_{LOW}} & if\ IMB(t_m)\ <\ 5\ U \\ f_{3_{LOW}}\ +\ \frac{1}{3}\Delta f_3 & if\ 5 \le IMB(t_m)\ <\ 10\ U \\ f_{3_{LOW}}\ +\ \frac{2}{3}\Delta f_3 & if\ 10 \le IMB(t_m)\ <\ 15\ U \\ f_{3_{HIGH}} & if\ IMB(t_m)\ \ge\ 15\ U \end{cases}$$

where $\Delta f_i = \left(f_{i_{HIGH}} - f_{i_{LOW}}\right)_{i\,=\,1,\,3}$. Intuitively, if C_2 is predicted, no adjustment of IMB is performed. Moreover, by multiplying the modulation function by $p(C_i)$, the adjustment of IMB became directly proportional to how much XGB was "certain" about assigning the current input to class C_i. This makes the adjustment robust to outliers, since in these cases, $p(C_i)$ results smaller, along with the IMB increment/decrement. Finally, the choice of using different adjustment depending on the original IMB amount allows the avoidance of too conservative, or too aggressive, adjustments when IMB is small, or big.

3.2. Simulated Scenario

XGB-IMB was tested in silico on a new dataset generated from the 100 virtual adult subjects of T1D-DM. Note that such a dataset is different from the one used to train and test the XGB model. In particular, two scenarios were created. In scenario A, 20 out of 100 subjects were used to tune the four XGB-IMB parameters ($f_{1_{LOW}}$, $f_{1_{HIGH}}$, $f_{3_{LOW}}$, and $f_{3_{HIGH}}$) over one-week simulation. Specifically, the parameter set was chosen by using a grid-search strategy, which uses a cost function that minimizes the average blood glucose risk index (BGRI) [30] of the studied population. In scenario B, the remaining 80 subjects were evaluated over one-month simulation with the aim of evaluating the performance of XGB-IMB strategy (with parameters obtained from scenario A). Finally, a comparison with the standard bolus calculator formula [31] was performed (SF-IMB).

3.3. Assessment of Glycemic Outcomes

For each subject of scenario B, performance of XGB-IMB and SF-IMB were compared in terms of: Mean (MEAN$_{BG}$) and standard deviation of glucose concentration (SD$_{BG}$); BGRI, percentage time in hypoglycemia (<70 mg/dL; %T$_{HYPO}$); percentage time in hyperglycemia (>180 mg/dL; %T$_{HYPER}$);

percentage time in glucose target range [70, 180] mg/dL (%T$_{TARGET}$); and percentage time in tight glucose target range [90, 140] mg/dL (%T$_{TTARGET}$). These metrics are commonly used to assess glycemic outcomes in T1D and follow a recent consensus report on outcome measures for artificial pancreas clinical trials evaluation [32].

These population metrics are reported as mean (± standard deviation) for Gaussian distributed metrics or median [interquartile range]. For this purpose, non-Gaussianity of each distribution was checked by mean of the Lilliefors test with a 5% confidence level. Finally, assessment of statistical significance between-method differences was performed, with a 1% confidence level, using a paired *t*-test or the Wilcoxon rank sum test if the compared distributions were both Gaussian or not, respectively.

The four XGB-IMB parameters obtained from scenario A are $f_{1LOW} = -0.3, f_{1HIGH} = -0.1, f_{3LOW} = 0.8$, and $f_{3HIGH} = 0.5$. Note that, while XGB-IMB reduces IMB up to 30%/10% if future hypoglycemia (C$_1$) is detected, it is quite aggressive at increasing IMB if future hyperglycemia is predicted (C$_3$), i.e., up to 80%/50% if the original IMB is smaller/bigger than 5/15 units.

Table 1 shows the population results obtained in the 80 subjects and corresponding to scenario B. Overall, XGB-IMB improves glycemic control across the population. In particular, when comparing SF-IMB vs. XGB-IMB, a statistically significant ($p < 0.01$) reduction of the overall glucose control in terms of MEAN$_{BG}$ and SD$_{BG}$ was achieved, i.e., 6.07 mg/dL and 4.95 mg/dL, respectively. Additionally, XGB-IMB lowered the glycemic risk in terms of BGRI ($p > 0.01$) by 1.36. Consistently, percentage time metrics showed better results when XGB-IMB was used. In detail, a significant improvement ($p < 0.01$) in %T$_{HYPER}$ (5.34%, from 35.18% to 29.84%), %T$_{TARGET}$ (5.02%, from 61.98% to 67.00%), and %T$_{TTARGET}$ (2.95%, from 28.22% to 31.17%) was observed, while a non-significant improvement ($p = 0.34$) of %T$_{HYPO}$ (0.11%, from 1.93% to 1.82%) was observed.

Table 1. Obtained on 80 virtual adult subjects on Scenario B.

	SF-IMB	XGB-IMB	*P*-VALUE
MEAN$_{BG}$	167.12 [155.28, 181.16]	161.05 [151.74, 169.33]	<0.01 *
SD$_{BG}$	54.97 [48.03, 63.95]	51.02 [44.92, 59.97]	<0.01 *
BGRI	9.36 [7.06, 11.43]	8.00 [6.60, 9.83]	<0.01 *
%T$_{HYPO}$	1.93 [0.07, 3.81]	1.82 [0.09, 3.81]	0.34
%T$_{HYPER}$	35.18 (±14.06)	29.84 (±11.50)	<0.01 **
%T$_{TARGET}$	61.98 (±13.89)	67.00 (±11.54)	<0.01 **
%T$_{TTARGET}$	28.22 [18.54, 40.56]	31.17 [24.49, 42.60]	<0.01 *

Results obtained using SF-IMB and XGB-IMB. Median [interquartile range] is reported for MEAN$_{BG}$, SD$_{BG}$, BGRI, %T$_{HYPO}$, %T$_{TTARGET}$; mean (± standard deviation) are reported for %T$_{HYPER}$, %T$_{TARGET}$. *: Statistically significant difference between XGB-IMB and SF-IMB using the Wilcoxon rank sum test. **: Statistically significant difference between XGB-IMB and SF-IMB using the *t*-test.

4. Discussion and Conclusions

In type 1 diabetes (T1D) management, continuous glucose monitors, insulin pumps, and activity monitors, together with additional physiological data streams (e.g., carbohydrate counting) can be used to develop new algorithms aimed at improving standard insulin therapy [33,34]. In this work, a new methodology based on a state-of-the-art machine learning model, i.e., the extreme gradient-boosted tree (XGB) model, is used to predict at meal-time postprandial glycemic status. Preliminary results obtained using in silico data generated with a state-of-the-art T1D simulator, show that XGB is accurate at discriminating between the selected post-prandial glycemic classes (i.e., hypoglycemia, hyperglycemia, and euglycemia).

The XGB model outcome can potentially be used for different purposes (e.g., glucose alerts, decision support on insulin therapy). Here, we investigated how the prediction for the postprandial glycemic status at meal-time can be used to adjust the meal insulin bolus dose. Results obtained through a simple set of heuristic rules to adjust the meal insulin bolus, confirm that the proposed technique has the potential to improve post-prandial glycemic control in a T1D population. From a practical point of view, the proposed XGB algorithm could be easily integrated in currently available insulin pumps, or implemented in a stand-alone mobile application.

The presented study has some limitations that need to be addressed in future work. Machine learning techniques assume that data is independent and identically distributed. Thus, given the finite nature of the dataset used to train the model, wrong predictions may occur due to the approximations done at representing the underneath input–output correlation. To solve this issue, safety constraints need to be devised to improve the robustness of the XGB model and avoid misclassification errors. Another important aspect that needs to be further studied is the feature selection process. Given the inherent capability of XGB at automatically discriminating important features, no preliminary feature selection was performed during the data preparation phase. However, future work is needed to investigate if it is possible to reduce its dimensionality and improve the interpretability of the results. Furthermore, additional work can be done to develop a more specific and optimized policy to adjust the meal insulin bolus according to XGB classifications. Finally, although the obtained results using a state-of-the-art in silico environment are very encouraging, these are still preliminary results. Therefore, the next logical step will certainly involve the validation of the proposed technique using retrospective clinical data and its subsequent evaluation in a prospective clinical trial.

References

1. American Diabetes Association, Standards of medical care in diabetes, Diabetes Care 2017, 41, S1–S159. Available online: https://doi.org/10.2337/dc18-S002 (accessed on 6 June 2019).
2. World Health Organization (WHO), Definition, diagnosis and classification of diabetes mellitus and its complications: Report of a WHO consultation. Part 1, diagnosis and classification of diabetes mellitus, Geneva, 1999. Available online: http://www.who.int/iris/handle/10665/66040 (accessed on 6 June 2019).
3. Cappon, G.; Marturano, F.; Vettoretti, M.; Facchinetti, A.; Sparacino, G. In silico assessment of literature insulin bolus calculation methods accounting for glucose rate of change. *J. Diabetes Sci. Technol.* **2018**, *13*, 103–110. [CrossRef] [PubMed]
4. Lane, J.E.; Shivers, J.P.; Zisser, H. Continuous glucose monitors: Current status and future developments. *Curr. Opin. Endocrinol. Diabetes Obes.* **2013**, *20*, 106–111. [CrossRef]
5. Lodwig, V.; Kulzer, B.; Schnell, O.; Heinemann, L. Current trends in continuous glucose monitoring. *J. Diabetes Sci. Technol.* **2014**, *8*, 390–396. [CrossRef] [PubMed]
6. Cappon, G.; Acciaroli, G.; Vettoretti, M.; Facchinetti, A.; Sparacino, G. Wearable continuous glucose monitoring sensors: A revolution in diabetes treatment. *Electronics* **2017**, *6*, 65. [CrossRef]
7. Oviedo, S.; Vehi, J.; Calm, R.; Armengol, J. A review of personalized blood glucose prediction strategies for T1DM patients. *Int. J. Number Method Biomed. Eng.* **2017**, *33*, e2833. [CrossRef] [PubMed]
8. Cappon, G.; Vettoretti, M.; Facchinetti, A.; Sparacino, A. A neural-network-based approach to personalize insulin bolus calculation using continuous glucose monitoring. *J. Diabetes Sci. Technol.* **2018**, *12*, 265–272. [CrossRef]
9. Herrero, P.; Pesl, P.; Reddy, M.; Oliver, N.; Georgiou, P.; Toumazou, C. Advanced insulin bolus advisor based on run-to-run control and case-based reasoning. *IEEE J. Biomed. Health Inform.* **2015**, *19*, 1087–1096.
10. Gadeleta, M.; Facchinetti, A.; Grisan, E.; Rossi, M. Prediction of adverse events from continuous glucose monitoring signal. *IEEE J. Biomed. Health Inform.* **2018**, *23*, 650–659. [CrossRef]

11. Aiello, E.M.; Toffanin, C.; Messori, M.; Cobelli, C. Postprandial glucose regulation via KNN meal classification in type 1 diabetes. *IEEE Control Sys. Letters* **2019**, *3*, 230–235. [CrossRef]

12. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGDD Int Conf on Knowledge Discovery, San Francisco, CA, USA, August 2016; pp. 13–17.

13. Schiavon, M.; Dalla Man, C.; Cobelli, C. Insulin sensitivity index based optimization of insulin to carbohydrate ratio: In silico study shows efficacious protection against hypoglycemic events caused by suboptimal therapy. *Diabetes Technol. Ther.* **2018**, *20*, 98–105. [CrossRef]

14. Contreras, I.; Oviedo, S.; Vettoretti, M.; Visentin, R.; Vehí, J. Personalized blood glucose prediction: A hybrid approach using grammatical evolution and physiological models. *PLoS ONE* **2017**, *12*, 1–16. [CrossRef] [PubMed]

15. Toffanin, C.; Visentin, R.; Messori, M.; Di Palma, F.; Magni, L.; Cobelli, C. Towards a run-to-run adaptive artificial pancreas: In silico results. *IEEE Trans. Biomed. Eng.* **2017**, *65*, 479–488. [CrossRef] [PubMed]

16. Vettoretti, M.; Facchinetti, A.; Sparacino, G.; Cobelli, C. Type-1 diabetes patient decision simulator for in silico testing safety and effectiveness of insulin treatments". *IEEE Trans. Biomed. Eng.* **2018**, *65*, 1281–1290. [CrossRef] [PubMed]

17. Dalla Man, C.; Micheletto, F.; Lv, D.; Breton, M.D.; Kovatchev, B.; Cobelli, C. The UVa/Padova type 1 diabetes simulator: New features. *J. Diabetes Sci. Technol.* **2017**, *8*, 26–34.

18. Herrero, P.; Pesl, P.; Bondia, J.; Reddy, M.; Oliver, N.; Georgiou, P.; Toumazou, C. Method for automatic adjustment of an insulin bolus calculator: In silico robustness evaluation under intra-day variability. *Comput. Methods Programs Biomed.* **2015**, *119*, 1–8. [CrossRef] [PubMed]

19. Walsh, J.; Roberts, R.; Heinemann, L. Confusion regarding duraztion if insulin action: A potential source for major insulin dose errors by bolus calculators. *J. Diabetes Sci. Technol.* **2014**, *8*, 170–178. [CrossRef] [PubMed]

20. Guerra, S.; Sparacino, G.; Facchinetti, A.; Schiavon, M.; Dalla Man, C.; Cobelli, C. A dynamic risk measure from continuous glucose monitoring data. *Diabetes Technol. Ther.* **2011**, *13*, 843–852. [CrossRef]

21. Hovorka, R.; Canonico, V.; Chassin, L.J.; Haueter, U.; Massi-Benedetti, M.; Orsini Federici, M.; Pieber, T.R.; Schaller, H.C.; Schaupp, L.; Vering, T.; et al. Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. *Physiol. Meas.* **2004**, *25*, 905–920. [CrossRef]

22. Willinska, M.E.; Chassin, L.J.; Schaller, H.C.; Schaupp, L.; Pieber, T.R.; Hovorka, R. Insulin kinetics in type 1 diabetes: Continuous and bolus delivery of rapid acting insulin. *IEEE Trans. Biomed. Eng.* **2005**, *52*, 3–12. [CrossRef]

23. Howsmon, D.P.; Cameron, F.; Baysal, N.; Ly, T.T.; Forlenza, G.P.; Maahs, D.M.; Buckingham, B.A.; Hahn, J.; Bequette, B.W. Continuous glucose monitoring enables the detection of losses in infusion set actuation (LISAs). *Sensors* **2017**, *17*, 161. [CrossRef]

24. Trevor, J.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed.; Springer: New York, NY, USA, 2009; ISBN 978-0-387-84857-0.

25. Florkowski, C.M. Sensitivity, specificity, receiver-operating characteristic (ROC) curves and likelihood ratios: Communicating the performance of diagnostic tests. *Clin. Biochem. Rev.* **2008**, *29*, S83–S87.

26. Bergstra, J.; Bardenet, R.; Bengio, Y.; Kegl, B. Algorithms for hyper-parameter optimization. In Proceedings of the 24th Int Conf in Advances in Neural Information Processing Systems, Granada, Spain, December 2015; pp. 12–15.

27. Vettoretti, M.; Facchinetti, A.; Sparacino, G.; Cobelli, C. A model of self-monitoring blood glucose measurement error. *J. Diabetes Sci. Technol.* **2017**, *11*, 724–735. [CrossRef] [PubMed]

28. Facchinetti, A.; Del Favero, S.; Sparacino, G.; Cobelli, C. Model of glucose sensor error components: Identification and assessment for new Dexcom G4 generation devices. *Med. Biol. Eng. Comput.* **2015**, *53*, 1259–1269. [CrossRef]

29. Visentin, R.; Dalla Man, C.; Kudva, Y.C.; Basu, A.; Cobelli, C. Circadian variability of insulin sensitivity: Physiological input for in silico artificial pancreas. *Diabetes Technol. Ther.* **2015**, *17*, 1–7. [CrossRef]

30. Fabris, C.; Patek, S.D.; Breton, M.D. Are risk indices derived from CGM interchangeable with SMBG-based indices? *J. Diabetes Sci. Technol.* **2015**, *10*, 50–59. [CrossRef] [PubMed]

31. Schmidt, S. Bolus calculators. *J. Diabetes Sci. Technol.* **2014**, *8*, 1035–1041. [CrossRef]

32. Maahs, D.M.; Buckingham, B.A.; Castle, J.R.; Cinar, A.; Damiano, E.R.; Dassau, E.; DeVries, J.H.; Doyle, F.J., III; Griffen, S.C.; Haidar, A.; et al. Outcome measures for artificial pancreas clinical trials: A consensus report. *Diabetes Care* **2016**, *39*, 1175–1179. [CrossRef]

33. Vettoretti, M.; Cappon, G.; Acciaroli, G.; Facchinetti, A.; Sparacino, G. Continuous glucose monitoring: Current use in diabetes management and possible future applications. *J. Diabetes Sci. Technol.* **2018**, *12*, 1064–1071. [CrossRef]

34. Facchinetti, A. Continuous glucose monitoring sensors: Past, present and future algorithmic challenges. *Sensors* **2016**, *16*, 2093. [CrossRef]