SugerMate: Towards an Ubiquitous Blood Glucose Tracking for Daily Use

WEIXI GU, Tsinghua University

Abstract here

ACM Reference format:
Weixi Gu. 2017. SugerMate: Towards an Ubiquitous Blood Glucose Tracking for Daily Use. *PACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 0, Article 0 (June 2017), 18 pages.
DOI: 0000001.0000001

1 INTRODUCTION

Blood glucose concentration plays an important role in personal health. Hyperglycemia (high blood glucose level) results in diabetes, leading to health risks such as pancreatic function failure, immunity reduce and ocular fundus diseases [30]. Meanwhile, hypoglycemia (low blood glucose level) also brings complications such as confusion, shakiness, anxiety, and if not treated in time, coma or death [24]. The International Diabetes Federation (IDF) reports that there were 415 million diabetic patients in 2015 and the number will rise to 642 million by 2040 [21]. People with diabetes need tight control of their blood glucose concentration to avoid both short-term and long-term physiological complications. While non-diabetic people normally have adequate self-regulation of blood glucose concentration, they can still risk hypoglycemia when taking prolonged exercises, drinking excess amounts of alcohol, having eating disorders, taking certain medicines (*e.g.*, certain painkiller and antibiotic), or having pre-diabetes [19, 22].

While continuous or regular blood glucose monitoring is essential for blood glucose management and beneficial for hyper- and hypoglycemia warning, it can be invasive and inconvenient to track blood glucose, especially during daily life. A standard and direct blood glucose measurement is to collect and analyze a drop of blood by finger pricking, which requires a new prick on the finger for every new observation. Alternatively, non-invasive (without penetrating the skin) continuous glucose monitoring (CGM) has attracted extensive research leveraging techniques such as thermal infrared spectroscopy, Raman spectroscopy and impedance spectroscopy [17, 48]. However, most CGM devices are expensive, cumbersome to wear for extended time, requires complicated operation and are usually limited to clinical uses, making them unattractive for both diabetic patients and non-diabetic people.

Towards more ubiquitous blood glucose monitoring when traditional CGM devices are unavailable or inconvenient to wear, researchers propose to explore the increasingly rich sensors embedded in commercial fitness wearables and smartphones as a complement. In addition to the glucose metabolism that is difficult to measure directly, blood glucose also correlates to easily measurable physiological activities such as food and drug intake, energy expenditure, sleep quality and emotional states [28]. Pioneer works [36, 42, 46]

Authors address: XXX and XXX E-mail: XXX .
Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
© 2017 ACM. 2474-9567/2017/6-ART0 $15.00
DOI: 0000001.0000001

have proposed preliminary systems leveraging commodity bio-sensors (*e.g.*, ECG electrodes) and fitness wearables (*e.g.*, accelerometer and galvanic skin response sensors) to predict blood glucose concentrations and alarm abnormal blood glucose events. Nevertheless, they are validated with limited numbers of dialectic patients [42, 46] or still require complex multi-sensory platforms [36, 46].

In this work, we design SugerMate, the first personalized smartphone-based non-invasive blood glucose monitoring system that detects abnormal blood glucose events by jointly tracking mean and insulin intake, physical activity and sleep quality. When SugerMate detects an abnormal blood glucose event, it reminds the user to double-check by finger pricking or using clinical CGM devices. SugerMate exploits recent advances in smartphone-based automatic human activity recognition [32] and sleep quality measurement [25] to acquire external factors as input for abnormal blood glucose event detection, making it widely applicable in daily use. In addition, SugerMate considers both *generic* and *user-specific* correlations between blood glucose levels and the measurable external factors, which is largely overlooked in previous works.

However, it is challenging to learn effective, accurate and personalized blood glucose models in practice. While there have been general blood glucose models that characterizing universal trends between blood glucose concentration and various external factors [37], they have to be adjusted based on user-specific data to account for inter-user differences [15]. Yet it is often difficult to collect sufficient data to directly build up personalized models [34]. *(i)* A disposable glucose sensor embedded in the CGM device is only capable of a few days [1], and most users are unwilling to wear CGM devices frequently due to discomfort. *(ii)* Despite their importance, measurements of hyper- and hypoglycemia events are rare compared with normal blood glucose concentrations, making it difficult to accurately detection abnormal blood glucose events.

To take full advantages of the *sparse*, *imbalanced* measurements to build *personalized* blood glucose level models, we propose gsMTRNN, a multi-task deep Recurrent Neural Network (RNN) framework to efficiently extract general blood glucose level relevant features and preserve user-specific characteristics. Evaluations on the blood glucose dataset composed of 112 users lasting 7 months show that our novel gsMTRNN framework outperforms both generic learning (*i.e.*, ignoring inter-user differences) and personalized learning (due to lack of measurements).

The key contributions of this work are summarized as follows.

- To the best of our knowledge, SugerMate is the first smartphone-based personalized abnormal blood glucose event detection system that works without CGM data as input. It automatically collects exercise levels and sleep quality, together with manual records of food and drug intake, and infers the current blood glucose level of users.
- We design gsMTRNN, a multi-task deep RNN framework able to share sensory measurements in an information representation layer and a temporal dynamic deep learning layer, but preserve individual blood glucose characteristics in a personality learning layer. It tackles the typical sparsity and imbalance problems in datasets for blood glucose modeling and offers an opportunity to build personalized blood glucose models for the general public based on limited personal measurements.
- We conduct extensive evaluations on both diabetic patients (type 1 and type 2) non-diabetic people. Experimental results from a dataset covering 35 healthy people, 38 Type I and 39 Type II patients during more than a half of a year demonstrate that SugerMate yields the accuracy of 82.14%, and outperforms than general learning method and personalized learning method in precision and recall respectively

In the rest of this paper, we...

2 RELATED WORK

Research on modeling blood glucose concentrations or abnormal events dates back to the 1960s and continues to attract extensive research interest [37]. Physiological models [16, 27] mathematically formulate the whole process of glucose metabolism and are widely used for simulations and studies involving glucose regulation. One major drawback of physiological models is that it requires prior knowledge to adjust the physiological parameters. Alternatively, researchers propose to combine machine learning techniques with a minimal physiological model or directly correlating blood glucose levels with insulin, food intake and other inputs without physiological parameters. For instance, Plis *et al.* [39] apply a generic physiological model of blood glucose dynamics to extract features for support vector regression to predict blood glucose levels. Reymann *et al.* [43] replace the physiological model by an online simulator and bring blood glucose tracking on mobile platforms.

While physiological models and the underlying glucose metabolism dominate the dynamics of blood glucose, there are also secondary impacting factors that depends on lifestyle. Variations of food intake, exercises, sleep quality, heart rates, etc., can also lead to blood glucose dynamics, which are not captured by a universal physiological model [28]. Consequently, it is essential to monitor external lifestyle factors such as food, exercise, sleep quality as input to improve the accuracy of blood glucose concentration modeling. METABO [23] is a client-server architecture based system that records dietary, physical activity, medication and medical information for hypoglycaemic and hyperglycaemic event prediction. Marling *et al.* [34] improve hypoglycemia detection by combining CGM data with heart rate, galvanic skin response and skin temperature collected from a fitness band. However, these works all require CGM data as input, making them invasive and inconvenient for both patients and non-diabetic people.

Alternatively, there has been attempt at non-invasive blood glucose monitoring with pervasive wearable and mobile devices. Nguyen *et al.* [36] observe distinct patterns in ECG signals during hypoglycemia and hyperglycemia in type 1 diabetic patients. Sobel *et al.* [46] integrate five types of sensory data from two accelerometers, a heat-flux sensor, a thermistor, two ECG electrodes and a galvanic skin response sensor to predict blood glucose concentration. Ranvier *et al.* [42] leverage ECG signals, and energy expenditure (estimated by an accelerometer and a breathing sensor) to detect hypoglycemic events. Our work is inspired by this category of research, and propose a smartphone-based non-invasive blood glucose monitoring system that jointly considers mean and insulin intake, physical activity and sleep quality without CGM data as input. Practically, we automatically record physical activity level and sleep quality without manual input, which notably improves the useability of our system.

Personalized blood glucose models are also important because models with generic parameters may not reflect user-specific factors such as age, weight and insulin-to-carbohydrates ratio [37]. Both the physiological parameters and the impact of life events on blood glucose need to be trained on user-specific data to account for inter-user differences [15]. However, a primary impediment is the lack of sufficient data to build up personalized models [34]. In this paper, we advance previous works by carefully designing a machine learning framework that shares blood glucose information among groups of users but preserves user-specific blood glucose characteristics via personalized learning, thus making full use of the limited, sometimes incomplete user-specific data, and achieving higher prediction accuracy than both generic learning and personalized learning.

3 PRELIMINARY

This section briefly reviews blood glucose levels and the impacting factors.

3.1 Blood Glucose Levels

The body of a healthy individual regulates the blood glucose concentration within 4.4 mmol/L to 6.1 mmol/L [49]. The blood glucose can grow slightly to 7.8 mmol/L after eating and usually returns to the normal range afterwards. Persons with the blood glucose concentration above 7.8 mmol/L (hyperglycemia) for a prolonged period are at the risk of diabetes mellitus, and need hypoglycemic drugs or insulin injection. In contrast, if the blood glucose concentration drops to below 4.4 mmol/L, it is a sign of hypoglycemia. Since we are interested in detecting normal and abnormal blood glucose events, we divide blood glucose concentration into 4 blood glucose levels as shown in Table 1.

Table 1. Normal and abnormal blood glucose levels.

Blood Glucose Value (mmol/L)	**Glocose Level**	**Explanation**
(0, 4.4]	Level 1	Low Blood Glucose
(4.4, 6.1]	Level 2	Normal Level of Fasting Blood Glucose
(6.1, 7.8]	Level 3	Normal Level of Postprandial Blood Glucose
$(7.8, +\infty)$	Level 4	High Blood Glucose

3.2 Impact Factors of Blood Glucose

The complete physiology of blood glucose is complex and it is believed that the blood glucose concentration is affected by both *internal* and *external* factors. Internal factors include the self-management of blood glucose, which is specific for each individual. External factors include activities that directly input or catabolize blood glucose such as food, drug, or insulin intake, and physical activities such as exercises. The external factors are mainly composed of food intakes, the exercise, the sleep quality and drugs or insulin inputs [18]. The physiological mechanism of external factors on blood glucose are shown as follows.

Food intake. The carbohydrates of the food intake can be quickly absorbed by gut and turned into the blood glucose. A steep rising trend usually occurs after consuming high-carbohydrate meals. Naturally, high-carbohydrate foods (*e.g.*, white grain products, bread and cookies) contributes higher blood glucose levels than low/moderate-carbohydrate foods (*e.g.*, meat, vegetables). Hence, the food intake is an important external factor of blood glucose.

Exercise. Exercise impacts the blood glucose level by the energy consumption. The movement of muscles will trigger the body cells absorbing sugar from blood for energy of doing exercise. By doing exercises, physical activity can help lower the blood glucose level for several hours after stop moving. The regular exercises promote the sensitivity of cells to the insulin, which can help keep blood glucose level vary in a normal range.

Sleep Quality. Sleep quality is strongly linked to the blood glucose levels. Several clinical studies [44, 45, 47] have demonstrates poor sleep quality is likely to results in high glucose levels. The body's reaction to sleep loss can resemble insulin resistance, a precursor to diabetes. Under this case, the cells of body easily fail to generate hormone efficiently, resulting in high blood glucose therefore. Moreover, sleep disorder is a key culprit of overweight [31, 41], which is also closed associated with the diabetes.

hypoglycemic drugs. The hypoglycemic drugs are mainly used for the patients with type II diabetes[20, 29, 38]. Usually, the anti-diabetes medication control the blood glucose concentrations by three ways: (1) stimulating the insulin secretion from pancreas, (2) promoting the tissue sensitivity to insulin, (3) low the rate at which glucose is absorbed from the gastrointestinal tract.

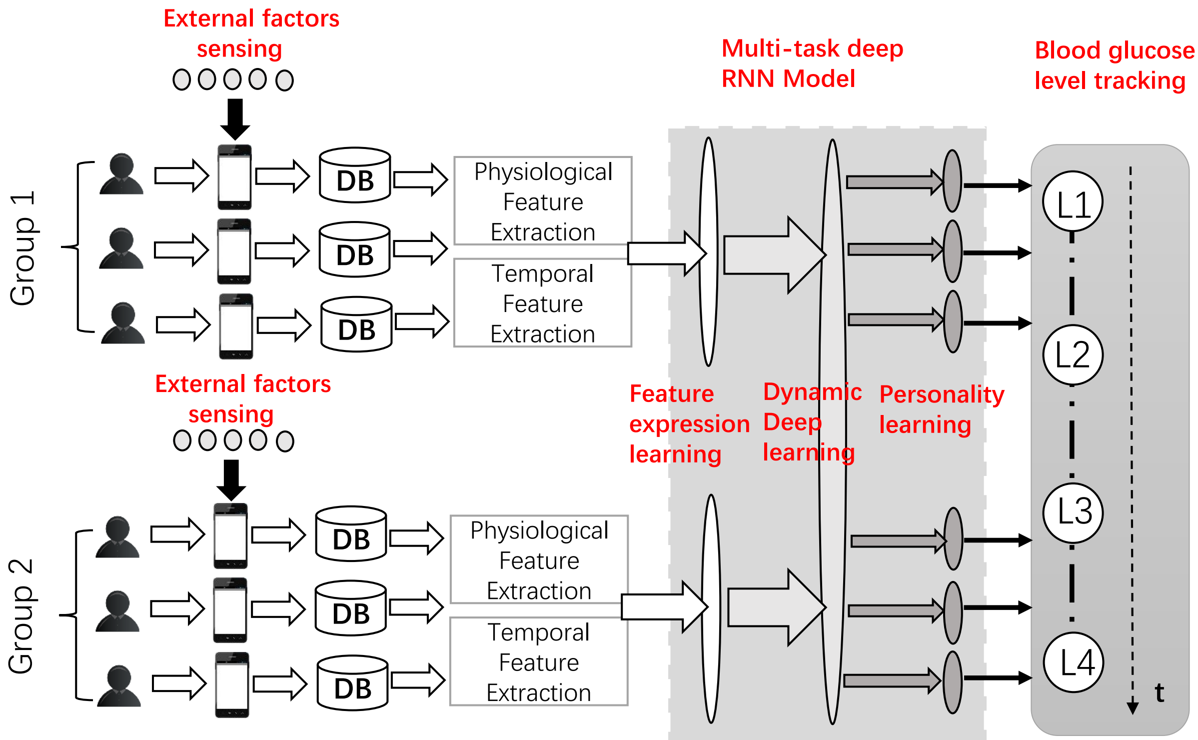

Fig. 1. The system architecture

Insulin. Insulin injection is a common way to treat the diabetes. For those who suffer Type I diabetes, they have to inject insulin every day due to their bodies fail to release insulin. For those who have Type II diabetes, insulin injection is also required when your body cannot proper react to insulin or produce enough of it by only taking drugs. By the assistance of insulin injection, the body can maintains the sugar circulating in the bloodstream within a healthy level.

Different to the internal factors, the impact trend of these external factors are similar to most people. And the blood glucose concentrations can be well modified by regulating these external factors.

4 OVERVIEW

The framework of SugerMate is composed of three major components: external factor collection module, multi-task deep RNN module and blood glucose level tracking module.

In the external factors collection module, the users are required to recrd their information into application, including of the diabetes type, drug, insulin and food intake. Meanwhile, SugerMate triggers the embedded sensors to sense the physical activities and sleep quality of user. SugerMate classifies the users into three groups based on glucose types afterwards. In the multi-task deep RNN (gsMTRNN) model, the feature representation is firstly learnt within the users in a same group. Then, a deep RNN layer is trained based on the dataset of all users, which is to establish the dynamic relationships between the outer contextual factors and the corresponding blood glucose level. In the last, gsMTRNN learns the personality of each user by the personality layer. Based on the results of gsMTRNN model, SugerMate tracks the current blood glucose level in the last module. Once SugerMate detects the abnormal points of blood glucose (*i.e.*, in a *high* or *low* blood glucose level), it reminds the user to measure the blood glucose by a clinical CGM or finger pricking method for a double-check.

The extraction mechanisms of outer contextual factors are detailed as follows.

Physical activity: SugerMate leverages the accelerometer to detect the user's activities by the approaches in [11], as well as the corresponding time costs. SugerMate then measures the calorie of user's physical consumption.

Food intake: SugerMate measures the food's effect on a person's blood glucose level based on the glycemic index.

Clinical drug intake: SugerMate records the name and amount of the drug that user eat.

Time: SugerMate invokes the timer embedded in the smartphone to record the time.

Sleep quality: SugerMate measures the user's sleep by the approach in [26].

5 DESIGN

5.1 External factor collection

Sensing module is mainly designed for collected the external factors of users.

Food intake. As the food intake is a main source of carbohydrate, SugerMate provides the food menu for users to record their daily intakes based on the carbohydrate food list [4]. Five common food categories have been provided by SugerMate, including grains ,vegetables, mike and egg,fruits and meats. The users are asked to enter their food items and the corresponding amounts. SugerMate calculate the carbohydrate of a meal and measure its impact on blood glucose level.

Drug intake. The oral diabetes medications enhance secretion of insulin into the blood by the pancreas or decrease amount of glucose released from liver, keeping the blood glucose in a low level for type II diabetes. In SugerMate, a drug menu of 11 oral diabetes is report for users to input their drug intake. After eating the diabetes drugs, users select their pills name and report the drug dosage. The drug list is provided based on [7]. SugerMate transfers the drug dose as the blood glucose efforts according to their work functions [7, 12, 13] by physiological model.

Insulin injection. Insulin injection is to control blood glucose concentration of those who have type I diabetes, and the patients of type II, whose blood sugar is too high for their bodies to control. SugerMate provides a insulin type list based on [5] for user with diabetes to enter their usage and insulin dosage, and then transfer it into physiological model to compute the blood glucose level.

Activity factors. Since the carbohydrate in the body can be consumed by daily exercise, resulting to varying the blood glucose level, SugerMate adopts an effective and power efficient approach [32] to automatically recognize six common user's daily activities (i.e., walking, running, upstairs, downstairs, sitting and standing), and record the corresponding durations. The caloric expenditure can be easily consumed by the calorie burn calculator formulas as Equation 1.

$$CalorieBurn = (BMR/24) * MET * T, \tag{1}$$

where BMR (Basal Methobolic Rate) is the amount of energy required to simply sit or lie quietly [35], and MET (Metabolic Equivalent) is the ratio of the work metabolic rate to the resting metabolic rate[9]. The calculator formula has been widely used by multiple sport applications [2, 3, 8]. SugerMate finally leverages the calories to measure the effects of exercise on the blood glucose.

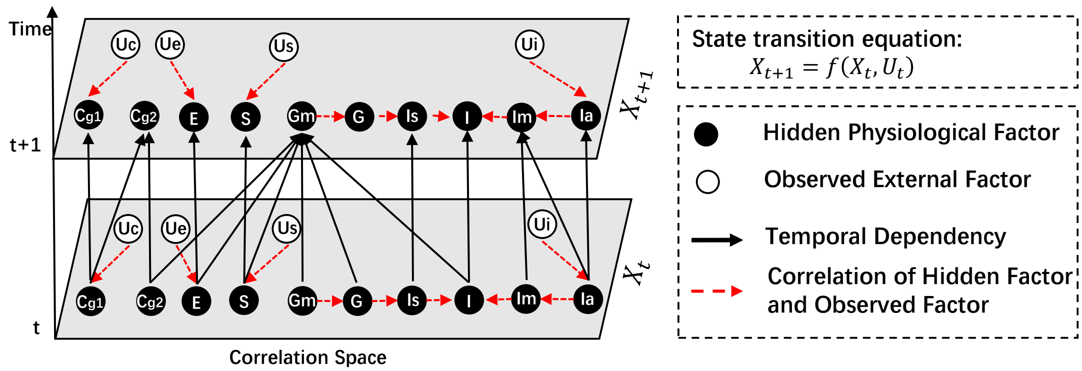

Fig. 2. Temporal graph of the physiological model.

Sleep quality. Sleep quality has a long influence on blood glucose level. To measure the sleep quality of users, SugerMate applies an effective method in [25] to measure the sleep quality, and leverages the sleep quality score as an index to evaluate the sleeping impact on blood glucose level.

5.2 Feature Engineering of Physiological-Temporal Views

5.2.1 Physiological View: F_p. The feature engineering of physiological view is to leverage a physiological model to quantify the dynamics of physiological factors in the body. The physiological model, based on the physiology mechanism of blood glucose in Section 3, has been widely studied in previous works [14, 18, 40].It primarily measures the real-time values of carbohydrate, insulin and glucose influenced by the external factors. We constructed our physiological model based on the work [18], with an extension of sleeping fact according to its physiological impact discussed in Section 3.

Temporal Graph of Physiological Model. The physiological model of SugerMate describes the physiological factors from five aspects: carbohydrate dynamics, insulin dynamics, exercise dynamics, sleep dynamics and blood glucose dynamics.

Carbohydrate dynamics: Carbohydrate dynamics refers to the transitions of carbohydrate consumption C_{g1} and the carbohydrate digestion C_{g2}. Equation 2 and Equation 3 show their transition equations respectively, where U_c stands for the carbohydrate proportion of meals.

$$C_{g1}(t+1) = C_{g1}(t) - \alpha_1^c * C_{g1}(t) + U_c(t) \tag{2}$$

$$C_{g2}(t+1) = C_{g2}(t) + \alpha_1^c * C_{g1}(t) - \alpha_2^c * C_{g2}(t) \tag{3}$$

Insulin dynamics: Insulin dynamics indicates the transitions of subcutaneous insulin absorption I_a (Equation 4), the insulin secretion by pancreas I_s and the insulin mass I_m (Equation 6). The level of active plasma insulin I (Equation 7). U_I states for the amount of insulin injected or simulated by the diabetes drugs. S^I and bm refer to the insulin sensitive and body mass respectively.

$$I_a(t+1) = I_a(t) - \alpha_{f,r,m}^I * I_a(t) + U_i(t) \tag{4}$$

$$I_s(t+1) = \begin{cases} I_s(t) + min[\alpha_1^I(\alpha_2^I(G_t - G_0)) + \alpha_3^I * G_0, \triangle I_{max}^s] & \text{Type II} \\ I_s(t) + 0 & \text{Type I} \end{cases} \tag{5}$$

$$I_m(t+1) = I_m(t) + \alpha^I_{f,r,m} * I_a(t) + \alpha^I_a * I_a(t) - \alpha^I_c * I_m(t) \tag{6}$$

$$I(t) = \frac{I_m(t) * S^I}{142 * bm} \tag{7}$$

Exercise dynamics: Exercise dynamics E denotes the exercise effect on insulin over the past time window. This long-term influence can be expressed by a cumulative moving average [6, 33] as Equation 8 in the physiological model.

$$E(t - t_0 + 1) = (t - t_0) * E(t - t_0) + U_e(t - t_0) \tag{8}$$

where t and t_0 are the current and beginning time point in the past time window. In SugerMate, the window size of exercise is set to 24 hours, which optimizes the experimental results and matches the conclusion of clinical studies.

Sleep dynamics: Sleep dynamics S represents the sleeping quality effect on insulin. In physiological model, sleeping effect also has a long-term influence on the insulin as the exercise effects. Specifically, it maintains a constant effect on blood glucose for each day. Equation 9 shows its transition equation.

$$S(t - t_0 + 1) = (t - t_0) * S(t - t_0) + U_s(t - t_0) \tag{9}$$

,

where t and t_0 are the current and beginning time point in the past time window. In SugerMate, the window size of sleep lasts for 7 days, which optimizes the experimental results and matches the conclusion of clinical studies.

Blood glucose dynamics: Blood glucose dynamic points to the fluctuation of glucose mass G_m in Equation 10, and glucose concentration $G(t)$ in Equation 11.

$$G_m(t+1) = G_m(t) + \delta_{abs} + \delta_{egp} - \delta_{ind} - \delta_{dep} - \delta_{clr} \tag{10}$$

$$G(t) = G_m(t)/(2.2 * bm) \tag{11}$$

Specifically, δ_{abs} in Equation 12 refers to the impact results of carbohydrate absorption, and δ_{egp} in Equation 13 indicates to the hepatic glucose production from the liver. These two factors improve the blood glucose concentration of body.

$$\delta_{abs} = \alpha^c_3 * \alpha^c_2 * C_{g2} \tag{12}$$

$$\delta_{egp} = \alpha^{egp}_2 * exp(-I(t)/\alpha^{egp}_3) - \alpha^{egp}_1 * G(t) \tag{13}$$

δ_{ind} (Equation 14) describes results of insulin independent uptake, which is consumed by the central nervous system and the red blood cells. δ_{dep} (Equation 15) indicates the impact results of insulin dependent uptake. It reflects the effects of insulin promoting muscle cells and fat cells to absorb glucose, which combines the influence of sleeping and exercise factors. δ_{clr} (Equation 16) stands for the influence of renal clearance on blood glucose. Once blood glucose concentrations exceeds the renal clearance threshold τ, the kidneys begin to remove excess glucose from the blood. The three factors decrease the blood glucose level.

$$\delta_{ind} = \alpha^{ind}_1 / \sqrt{G(t)} \tag{14}$$

$$\delta_{dep} = \alpha^{dep}_1 * E(t) * S(t) * I(t)/(G(t) + \alpha^{dep}_2) \tag{15}$$

$$\delta_{clr} = \alpha_1^{clr} * (G(t) - \tau) \tag{16}$$

The default values of parameters in the physiological model are set based on [18]. They are further tuned for each person by the classifier on the training dataset by 10-cross validations, which optimize the experimental results.

In SugerMate, we use smartphone to collect the external factors $U_t = \{U_c(t), U_e(t), U_s(t), U_i(t)\}$, and apply the physiological model to generate real-time observed vector $X_t = \{C_{g1}(t), C_{g2}(t), I_s(t), I_m(t), I_a(t), I(t), E(t), S(t), G_m(t), G(t)\}$ for every blood glucose sample at the corresponding time t. The hidden physiological factors at $t+1$ can be calculated by $X_{t+1} = f(X_t, U_t)$, where f is the station transition functions. SugerMate computes X at each time step and treats it as 10-dimensional features.

5.3 Temporal View: F_t

The blood glucose dynamic holds a natural temporal ordering, SugerMate mines its timing characteristics to predict the blood glucose level.

Three dimensional features of temporal view have been considered.

Historical Blood glucose trend: F_{t1}. Since the trend of blood glucose of a single person hardly occurs significant alteration within a short period, SugerMate calculates the historical blood glucose trend G_{Trend} by average the true blood glucose concentrations at each corresponding time stamp t over recent D days in the past by Equation 17.

$$G_{Trend} = \frac{1}{D}\sum_{d=1}^{D} G(f_t(d)), t = 1, 2, ..., N \tag{17}$$

where $N = 480$ refers to the number of time stamps in a day, and D reflects the days of measurement. $G(f_t(d))$ indicates the true blood glucose value measured by the CGM at the time stamp t of the dth day. In SugerMate, we select $D = 5$ based on the optimal experimental results

SugerMate treats it as a temporal feature to reflect the historical blood glucose trend over the same period.

Blood glucose value with similar physiological features: F_{t2}. As the blood glucose concentrate is determined by the factors of carbohydrate, insulin, exercise intensity and sleep quality, the similar blood glucose values should maintain similar values of these physiological factors. Accordingly, SugerMate applies the k-Nearest Neighbors algorithm [10] to search for 5 blood glucose values with the most similar physiological features, and average them as one dimensional feature.

Recent physiological features: F_{t3}. Considering the physiological features may generate the temporal delay effects on the blood glucose, SugerMate also takes the physiological factor vectors in the past 15 minutes as recent physiological features to measure their impacts on the current blood glucose concentration.

5.4 Blood glucose level prediction

Multi-division deep-dynamic RNN (Md^{3}RNN)

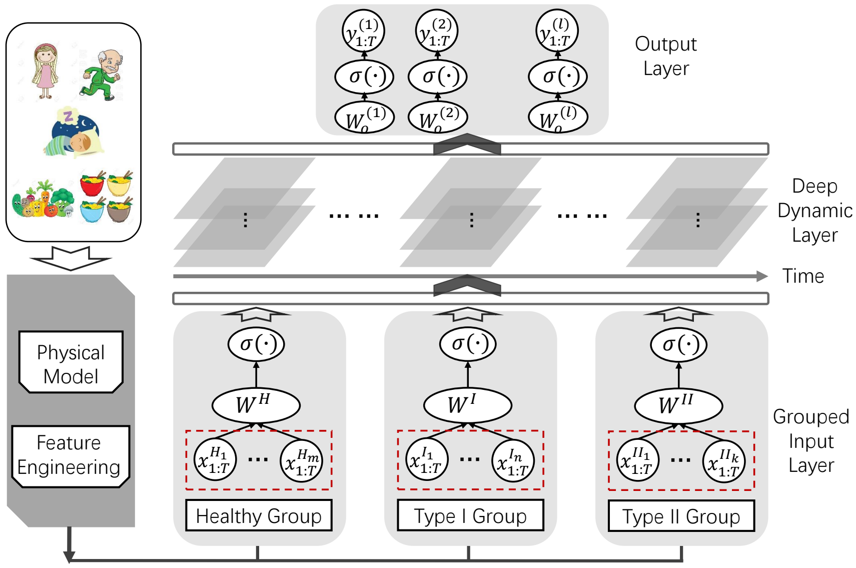

Fig. 3. The Md^{3}RNN structure

6 EVALUATION

6.1 Experimental Settings

Datasets. We validate SugerMate on a dataset of 112 participants (35 non-diabetes, 38 type I diabetic patients, 39 type II diabetic patients) collected during July 2016 to January 2017. Each participant is equipped with a **TODO: concrete brand, type** CGM device to record blood glucose concentration every 3 minutes and a smartphone with SugerMate installed to collect external factor information either automatically (activities and sleep quality) or manually (food, drug, and insulin intake). All participants agree to take measurements (*i.e.*, wear the CGM device and use SugerMate to record external factors) for at least 5 days. In total we obtain **TODO: XXX** samples of blood glucose concentration and the corresponding external factors covering **TODO: XXX** hours. In brief, we collect the following categories of data:

- **Personal information.** We record basic personal data including gender, age, weight and health status as reference for user grouping. Table 2 summarizes the basic information of the participants.
- **Blood glucose measurements.** Since we mainly aim to detect abnormal blood glucose events, we divide the blood glucose concentration into 4 levels: **TODO: the concrete range of each level** . The duration of blood glucose measurements varies from 5 to 30 days. Table 3 summarizes amount of blood glucose measurements.
- **External factor measurements.** During measurements of blood glucose concentration, each participant manually inputs the times of their daily meal, drug and insulin intake. SugerMate automatically records activity levels and sleep quality as in Sec. 5.1. **TODO: show user interface?**

Ground Truth. We use the blood glucose concentrations collected by the CGM device as ground truth.

Table 2. Summary of participant information.

(a)

Age (year)	# User
15-24	8
25-34	17
35-44	24
45-54	29
55-70	34

(b)

Weight (kg)	# User
30-44	18
45-54	21
55-64	32
65-74	22
75-90	19

(c)

Status	# User
Non-diabetes	35
Type I	38
Type II	39

(d)

Gender	# User
Male	57
Female	55

Table 3. Summary of blood glucose measurements.

(a)

Duration (days)	# User
6-10	48
11-15	24
16-20	20
21-25	13
26-30	7

(b)

Blood Glucose	# Sample
Level 1	75369
Level 2	293530
Level 3	235686
Level 4	158054
Total	762639

Metrics. We mainly adopt three metrics to evaluate the performance of SugerMate, including precision [], recall [] and accuracy [].

6.2 Overall Accuracy

Since all participants collected both measurements of CGM and external factors for at least 5 days, we use measurements during the former 4 days for training and the rest for testing. Table 4 shows the overall performance of SugerMate. All results are averaged over the testing data. As shown, the recalls and the precisions for all the 4 blood glucose levels are above 79% and 73%, respectively. In particular, the recalls for Level 1 (**TODO: low/high blood glucose?**) and Level 4 **TODO: low/high blood glucose?** are 83.13% and 85.23%, even though the training data for Level 1 and Level 4 only account for **TODO: XXX%** and **TODO: XXX%** of the entire training set. Overall, SugerMate yields an accuracy of 82.14%.

6.2.1 Effectiveness of Features. Table 5 shows the average precisions and recalls for all the 4 blood glucose levels with different combinations of features. By combining physiological features (F_p) with temporal features (F_t), the overall precisions and recalls improve by **TODO: XXX% to XXX%** .

Table 4. Confusion matrix of SugerMate.

Ground Truth	Predictions: Level 1	Level 2	Level 3	Level 4	Recall
Level 1	62657	5521	3672	3519	83.13%
Level 2	16346	240584	27563	9037	81.96%
Level 3	2660	30905	188472	13649	79.97%
Level 4	3443	5620	14278	134713	85.23%
Precision	73.62%	85.12%	80.55%	83.72%	Accuracy: 82.14%

TODO: XXX feature brings in the most notable improvement in detecting abnormal blood glucose events (Level XXX and Level XXX). This is because XXX.

Table 5. Effectiveness of features.

	Level 1		Level 2		Level 3		Level 4	
Features	**Precision**	**Recall**	**Precision**	**Recall**	**Precision**	**Recall**	**Precision**	**Recall**
F_p	43.37%	32.82%	46.03%	39.10%	51.79%	48.95%	56.30%	43.49%
F_p+F_{t1}	51.97%	58.11%	60.42%	58.90%	63.35%	53.59%	69.82%	55.16%
F_p+F_{t1}+F_{t2}	64.60%	73.08%	69.87%	61.23%	74.33%	67.81%	76.64%	72.32%
F_p+F_{t1}+F_{t2}+F_{t3}	73.62%	83.13%	83.13%	83.13%	80.55%	79.97%	83.72%	85.23%

6.2.2 Effectiveness of Multi-task Framework. To demonstrate the effectiveness of the multi-task framework in making full use of the training dataset, we compare gsMTRNN with two other frameworks.

- *General Learning.* All the training data are directly fed into the model for training indifferently. General learning results in a *generic* model that assumes universal correlations between all inputs and the blood glucose levels.
- *Single Learning.* We train a different model for each individual participant by feeding his/her own measurements into the model. Single learning results in a *personalized* model without sharing data and learning knowledge from measurements of other participants.

Fig. 4 shows the overall precisions and recalls of our multi-task learning framework as well as general learning and single learning. As shown, our multi-task learning framework outperforms both general learning and single learning by **TODO: XXX%** in precision and **TODO: XXX%** in recall, respectively. General learning performs slightly better than single learning for Level 2 and 3 (normal blood glucose levels), partly because the correlations between the inputs and normal blood glucose levels are relatively consistent for most people, while single learning suffers from lack of training data as it only uses user-specific data. Conversely, single learning achieves higher precision and recall than general learning for Level 1 and 4 (abnormal blood glucose levels), partly because there are notable inter-person differences in the correlations between the inputs and abnormal blood glucose levels. That is, the reasons for abnormal blood glucose levels can vary from person to person. The multi-task framework combines the advantages of both general learning and single learning, which makes better use of the limited training data by sharing measurements among users, while preserving user-specific characteristics via the personal learning layer.

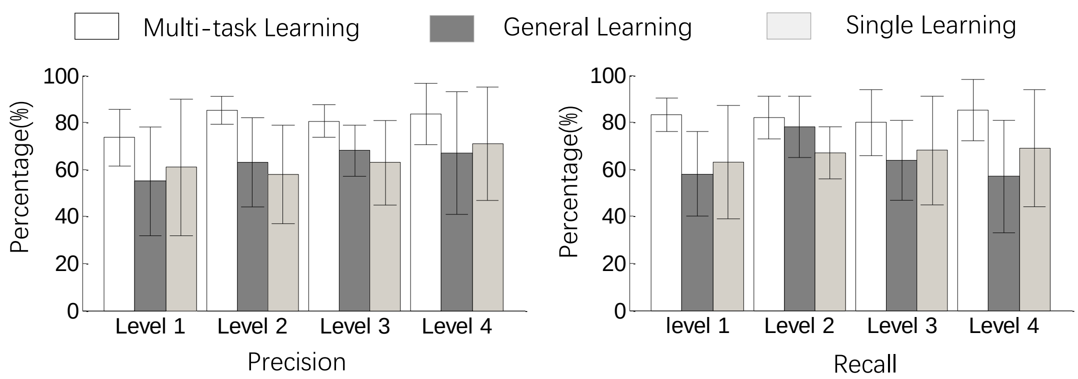

Fig. 4. Effectiveness of our multi-task learning framework in making full use of the training dataset.

6.2.3 Effectiveness of Deep Learning. To demonstrate the effectiveness of adopting deep learning algorithms over conventional shallow learning algorithms, we compare our gsMTRNN with the following mainstream learning algorithms. **TODO: to complete and add the corresponding references TODO: add a table summarizing the key parameters used for each model. Just mention the parameters for each model are optimized. No need to describe in such detail in the following.**

- **Gradient Boosting (GB).** GB generates a prediction model by combining many weak classifiers into a stronger classification committee. We use AdaBoost procedure implemented in the fastAdaboost package to combine basic tree classifiers for ensemble learning. We vary the maximum tree depth from 10 to 50 by factors of ten. The number of boosting iterations is varied from 100 to 500 by a step size of 50.
- **Support Vehicle Machine (SVM).** SVM bases on the idea of optimal separating hyperplane that maximizes the separation margin of two data groups (classes). Due to this construction, it usually generalizes well, and its dual form is a quadratic programing that can be easily incorporated with kernels. We train the Gaussian kernel SVM classifier with the kernlab package, which implements the sequential minimal optimization algorithm. We vary the kernel width from 2^{-5} to 2^4 with a factor of 2. We pick the penalty parameter from the set $\{10^i | i = -3, 0.5, 2\}$. To eliminate scale/location discrepancies among input variables, all features are normalized before being used in the training phase.
- **Hidden Markov model (HMM).**
- **Logistic Regression (LR).** LR models the posterior distribution of the class labels as a sigmoidal function of linear combinations of features. We use the glmnet package to train LR models with elastic net (combined L1 and L2) regularization. We vary the penalty parameter from 10^{-3} to 10^2 with a factor 5. The mixing parameter is varied from 0 (Ridge) to 1 (Lasso) by a step size of 0.1.
- **Random Forest (RF).** As another ensemble method, RF combines many simple decision trees together and output the mode of classes for prediction. To avoid correlation among base trees, random set of features are selected in the splitting process when constructing each decision tree. For implementation, we adopt the conditional inference tree algorithm in the Party package. The total number of trees is tuned from 100 to 1000, and the maximum tree depth from 10 to 50. The splitting threshold is also varied from 0.1 to 0.9 with 0.1 intervals for cross validation.

- **Gaussian Processes (GP).** Instead of directly parameterizing a latent function for classification, GP models it with a generic Gaussian process. The posterior of the process is updated with training data set, and is squashed through a logistic function for classification. We implement GP with the kernlab package, which includes several approximation algorithms for acceleration. We use the radial basis kernel and vary the kernel width from 2^{-5} to 2^{4} with an incremental factor of $2^{0.5}$.

TODO: results

6.3 Micro-benchmarks

In this section we evaluate the performance of SugerMate for different participant groups as well as its temporal performance.

6.3.1 Impact of amount of training samples. In this experiment, we evaluate the performance of SugerMate with increasing numbers of training samples. Since our dataset consists of measurements of durations from 5 to 30 days (see Table 3), we use measurements of 5 to 25 days for training and the remaining 5 days for testing. The results are averaged over all testing samples as in previous evaluations. Fig. 5 illustrates the results for all the 4 blood glucose levels. As expected, the precisions and recalls for all the 4 blood glucose levels improve smoothly with the increase of training samples. The results verify that the challenge (and our motivation to adopt a multi-task learning framework) is the lack of training data. Note that SugerMate is not a replacement of the current CGM devices, but rather, a complement when CGM devices are uncomfortable or inconvenient to wear. Therefore we envision the training dataset will grow gradually after multiple times of CGM device wearing, at least for diabetes patients, and the overall accuracy will also improve over time as a result.

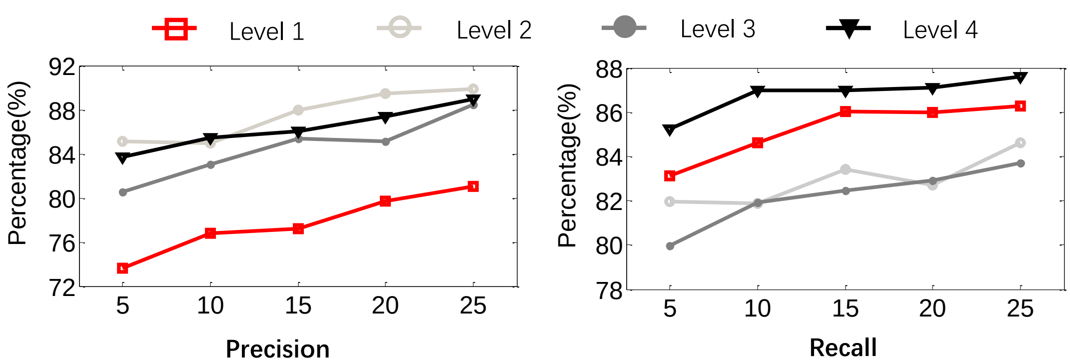

Fig. 5. Impact of increasing amount of training samples.

6.3.2 Impact of Temporal Gaps. The blood glucose concentration is correlated with the previous blood glucose concentration because of the control loop of the glucose metabolism [16, 27, 37]. Since SugerMate does not rely on the previous blood glucose level as an input, it is natural that the accuracy of SugerMate will degrade if there is a long gap between the training and the testing datasets (*i.e.*, the training dataset can be outdated). Fig. 6 plots the overall performance by training using the same 4 days of measurements, and testing on measurements collected on the 5-9th, 10-14th, 15-19th, 20-24th, and 25-29th days, respectively. As expected, both the precisions and recalls drop moderately with the increase of temporal gaps between the training and the testing datasets, with a maximum decrease of **TODO:**

XXX% and **TODO: XXX%** in precision and recall after XXX days. From the results, we recommend SugerMate users to put on the CGM device to **TODO: retrain the model?** at least every three weeks (?).

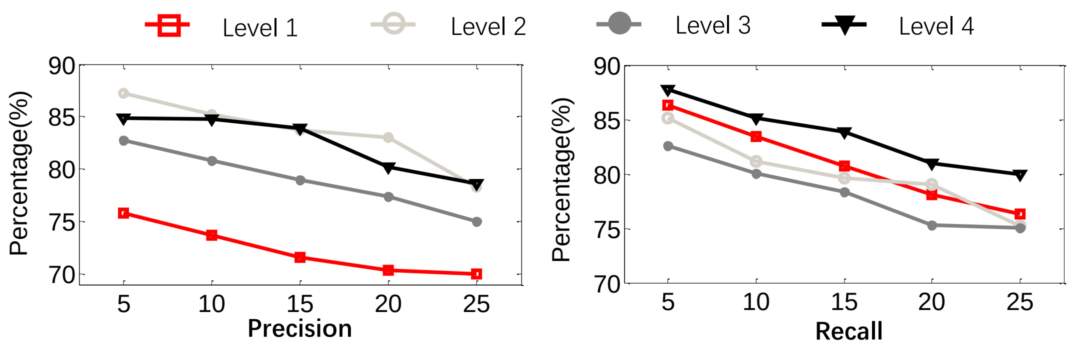

Fig. 6. Impact of temporal gaps between the training and testing datasets.

6.3.3 Detection Accuracy throughout a Day. Fig. 7 compares the predictive results of SugerMate and true blood glucose level of a user over one day. **TODO: redraw similar figures for a non-diabetic user, a type I user and a type II user** . **TODO: better to include a trace where the accuracy when taking exercises, taking drugs and foods, is high. Also need to explain what the user is doing when the prediction is wrong.**

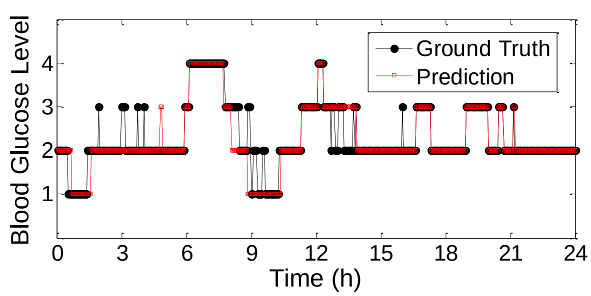

Fig. 7. The comparison of prediction and the ground truths of one user.

6.3.4 Detection Accuracy for Different Groups. **TODO: redraw the figures, no need to include general and personal learning here. show the performance for different health status, gender, age groups, weight groups, etc., explain why SugerMate works better for certain groups, or works well for all groups** .

We also apply general learning approach on the users in same group, and compare the prediction performance of gsMTRNN. Fig. 8 shows the results. As is shown, gsMTRNN outperforms the general learning methods in each group, especially the performance of level 1 and level 4. It mainly results by two reasons. On the one hand, the limitation of blood glucose data in each group weakens the capability of temporal dynamic characteristics. On the other hand, the imbalanced distribution of blood glucose

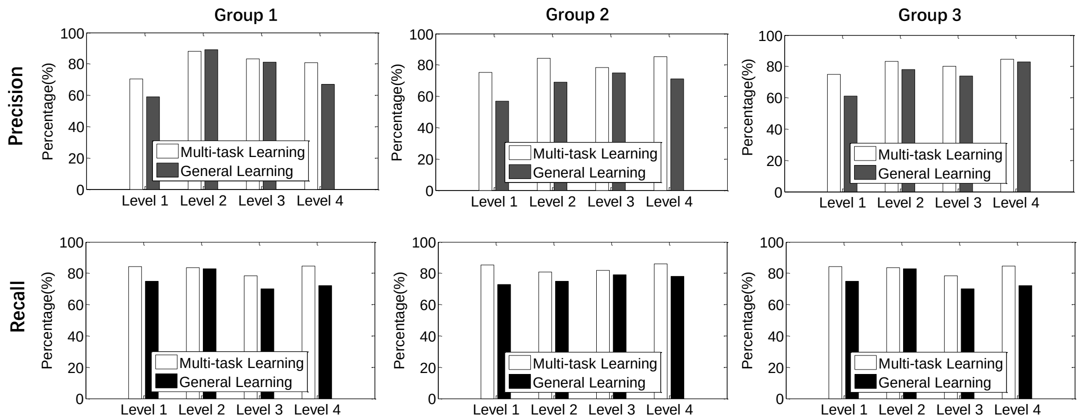

Fig. 8. The comparison of multi-task learning and group general learning.

data of one group also low the performance down. For example, much more data of level 4 and much less data of level 1 in group 3 (type II diabetes) low down the recall of level 1 and the precision of level 4. It is even hard to be solved by the cost sensitive approach.

7 CONCLUSION

REFERENCES

[1] Blood glucose monitoring. (????). https://en.wikipedia.org/wiki/Blood_glucose_monitoring.

[2] Calorie Counter - MyFitnessPal, note = https://play.google.com/store/apps/details?id=com.myfitnesspal.android&hl=en,. (????).

[3] Calories Burned Calculator, note = https://www.healthstatus.com/calculate/cbc,. (????).

[4] Diabetes: Carbohydrate Food List. (????). http://www.med.umich.edu/1libr/MEND/CarbList.pdf.

[5] Insulin A to Z: A Guide on Different Types of Insulin. (????). http://www.joslin.org/info/insulin_a_to_z_a_guide_on_different_types_of_insulin.html.

[6] Moving average, note = https://en.wikipedia.org/wiki/Moving_average,. (????).

[7] Oral Diabetes Medications Summary Chart. (????). http://www.joslin.org/info/oral_diabetes_medications_summary_chart.html.

[8] SHAPESENSE-INTELLIGENT FITNESS. (????). http://www.shapesense.com/.

[9] Barbara E Ainsworth, William L Haskell, Melicia C Whitt, Melinda L Irwin, Ann M Swartz, Scott J Strath, WILLIAM L O Brien, David R Bassett, Kathryn H Schmitz, Patricia O Emplaincourt, and others. 2000. Compendium of physical activities: an update of activity codes and MET intensities. *Medicine and science in sports and exercise* 32, 9; SUPP/1 (2000), S498–S504.

[10] Naomi S Altman. 1992. An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician* 46, 3 (1992), 175–185.

[11] Akram Bayat, Marc Pomplun, and Duc A Tran. 2014. A study on human activity recognition using accelerometer data from smartphones. *Procedia Computer Science* 34 (2014), 450–457.

[12] Wendy L Bennett, Lisa M Wilson, Shari Bolen, Nisa Maruthur, Sonal Singh, Ranee Chatterjee, Spyridon S Marinopoulos, Milo A Puhan, Padmini Ranasinghe, Wanda K Nicholson, and others. 2011. Oral diabetes medications for adults with type 2 diabetes: an update. (2011).

[13] Shari Bolen, Leonard Feldman, Jason Vassy, Lisa Wilson, Hsin-Chieh Yeh, Spyridon Marinopoulos, Crystal Wiley, Elizabeth Selvin, Renee Wilson, Eric B Bass, and others. 2007. Systematic review: comparative effectiveness and safety of oral medications for type 2 diabetes mellitus. *Annals of internal medicine* 147, 6 (2007), 386–399.

[14] Thomas Briegel and Volker Tresp. 2002. A Nonlinear State Space Model for the Blood Glucose Metabolism of a Diabetic (Ein nichtlineares Zustandsraummodell für den Blutglukosemetabolismus eines Diabetikers). *at-Automatisierungstechnik Methoden und Anwendungen der Steuerungs-, Regelungs-und Informationstechnik* 50, 5/2002 (2002), 228.

[15] Razvan Bunescu, Nigel Struble, Cindy Marling, Jay Shubrook, and Frank Schwartz. 2013. Blood glucose level prediction using physiological models and support vector regression. In *Proceedings of IEEE International Conference on Machine Learning and Applications*, Vol. 1. 135–140.
[16] Chiara Dalla Man, Robert A Rizza, and Claudio Cobelli. 2007. Meal simulation model of the glucose-insulin system. *Transactions on Biomedical Engineering* 54, 10 (2007), 1740–1749.
[17] Carlos Eduardo Ferrante do Amaral and Benhard Wolf. 2008. Current development in non-invasive glucose monitoring. *Medical engineering & physics* 30, 5 (2008), 541–549.
[18] David L Duke. 2010. *Intelligent diabetes assistant: A telemedicine system for modeling and managing blood glucose*. CARNEGIE MELLON UNIVERSITY.
[19] Margaret Eckert-Norton and Susan Kirk. 2013. Non-diabetic hypoglycemia. *The Journal of Clinical Endocrinology & Metabolism* 98, 10 (2013), 39A–40A.
[20] Dean T Eurich, Finlay A McAlister, David F Blackburn, Sumit R Majumdar, Ross T Tsuyuki, Janice Varney, and Jeffrey A Johnson. 2007. Benefits and harms of antidiabetic agents in patients with diabetes and heart failure: systematic review. *Bmj* 335, 7618 (2007), 497.
[21] International Diabetes Federation. 2015. IDF Diabetes Atlas. 7th ed. (2015). http://www.diabetesatlas.org/.
[22] Philip Felig, Ali Cherif, Akira Minagawa, and John Wahren. 1982. Hypoglycemia during prolonged exercise in normal men. *New England Journal of Medicine* 306, 15 (1982), 895–900.
[23] Eleni Georga, Vasilios Protopappas, Alejandra Guillen, Giuseppe Fico, Diego Ardigo, Maria Teresa Arredondo, Themis P Exarchos, Demosthenes Polyzos, and Dimitrios I Fotiadis. 2009. Data mining for blood glucose prediction and knowledge discovery in diabetic patients: The METABO diabetes modeling and management system. In *Proceedings of IEEE Annual International Conference of the Engineering in Medicine and Biology Society*. 5633–5636.
[24] DCCT Research Group and others. 1991. Epidemiology of severe hypoglycemia in the Diabetes Control and Complications Trial. *The American Journal of Medicine* 90, 4 (1991), 450–459.
[25] Weixi Gu, Zheng Yang, Longfei Shangguan, Wei Sun, Kun Jin, and Yunhao Liu. 2014. Intelligent sleep stage mining service with smartphones. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 649–660.
[26] Weixi Gu, Zheng Yang, Longfei Shangguan, Wei Sun, Kun Jin, and Yunhao Liu. 2014. Intelligent sleep stage mining service with smartphones. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 649–660.
[27] Roman Hovorka, Valentina Canonico, Ludovic J Chassin, Ulrich Haueter, Massimo Massi-Benedetti, Marco Orsini Federici, Thomas R Pieber, Helga C Schaller, Lukas Schaupp, Thomas Vering, and others. 2004. Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. *Physiological Measurement* 25, 4 (2004), 905.
[28] Shingo Iwasaki, Junji Kozawa, Kenji Fukui, Hiromi Iwahashi, Akihisa Imagawa, and Iichiro Shimomura. 2015. Coefficient of variation of R-R interval closely correlates with glycemic variability assessed by continuous glucose monitoring in insulin-depleted patients with type 1 diabetes. *Diabetes Research and Clinical Practice* 109, 2 (2015), 397–403.
[29] Mankil Jung, Moonsoo Park, Hyun C Lee, Yoon-Ho Kang, Eun S Kang, and Sang K Kim. 2006. Antidiabetic agents from medicinal plants. *Current medicinal chemistry* 13, 10 (2006), 1203–1218.
[30] Ronald Klein. 1995. Hyperglycemia and microvascular and macrovascular disease in diabetes. *Diabetes Care* 18, 2 (1995), 258–268.
[31] Kristen L Knutson, Armand M Ryden, Bryce A Mander, and Eve Van Cauter. 2006. Role of sleep duration and quality in the risk and severity of type 2 diabetes mellitus. *Archives of internal medicine* 166, 16 (2006), 1768–1774.
[32] Jennifer R Kwapisz, Gary M Weiss, and Samuel A Moore. 2011. Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter* 12, 2 (2011), 74–82.
[33] Cynthia A Lowry, William H Woodall, Charles W Champ, and Steven E Rigdon. 1992. A multivariate exponentially weighted moving average control chart. *Technometrics* 34, 1 (1992), 46–53.
[34] Cindy Marling, Lijie Xia, Razvan Bunescu, and Frank Schwartz. 2016. Machine Learning Experiments with Noninvasive Sensors for Hypoglycemia Detection. In *Proceedings of IJCAI Workshop on Knowledge Discovery in Healthcare Data*.
[35] Brian K McNab. 1997. On the utility of uniformity in the definition of basal rate of metabolism. *Physiological Zoology* 70, 6 (1997), 718–720.
[36] Linh Lan Nguyen, Steven Su, and Hung T Nguyen. 2012. Identification of hypoglycemia and hyperglycemia in type 1 diabetic patients using ECG parameters. In *Proceedings of IEEE Annual International Conference of the Engineering in Medicine and Biology Society*. 2716–2719.
[37] Silvia Oviedo, Josep Vehí, Remei Calm, and Joaquim Armengol. 2016. A review of personalized blood glucose prediction strategies for T1DM patients. *International Journal for Numerical Methods in Biomedical Engineering* (2016).

[38] DK Patel, SK Prasad, R Kumar, and S Hemalatha. 2012. An overview on antidiabetic medicinal plants having insulin mimetic property. *Asian Pacific journal of tropical biomedicine* 2, 4 (2012), 320–330.

[39] Kevin Plis, Razvan Bunescu, Cindy Marling, Jay Shubrook, and Frank Schwartz. 2014. A Machine Learning Approach to Predicting Blood Glucose Levels for Diabetes Management. In *Proceedings of AAAI Workshop on Modern Artificial Intelligence for Health Analytics*.

[40] Kevin Plis, Razvan Bunescu, Cindy Marling, Jay Shubrook, and Frank Schwartz. 2014. A machine learning approach to predicting blood glucose levels for diabetes management. *Modern Artificial Intelligence for Health Analytics. Papers from the AAAI-14* (2014).

[41] Naresh M Punjabi, John D Sorkin, Leslie I Katzel, Andrew P Goldberg, Alan R Schwartz, and Philip L Smith. 2002. Sleep-disordered breathing and insulin resistance in middle-aged and overweight men. *American journal of respiratory and critical care medicine* 165, 5 (2002), 677–682.

[42] Jean-Eudes Ranvier, Fabien Dubosson, Jean-Paul Calbimonte, and Karl Aberer. 2016. Detection of hypoglycemic events through wearable sensors. In *Proceedings of the International Workshop on Semantic Web Technologies for Mobile and PErvasive Environments*.

[43] Maximilian P Reymann, Eva Dorschky, Benjamin H Groh, Christine Martindale, Peter Blank, and Bjoern M Eskofier. 2016. Blood glucose level prediction based on support vector regression using mobile platforms. In *Proceedings of IEEE Annual International Conference of the Engineering in Medicine and Biology Society*. 2990–2993.

[44] ANDRE J Scheen, MARIA M Byrne, Laurence Plat, Rachel Leproult, and Eve Van Cauter. 1996. Relationships between sleep quality and glucose regulation in normal humans. *American Journal of Physiology-Endocrinology And Metabolism* 271, 2 (1996), E261–E270.

[45] Andr&eacute J Scheen and Eve Van Cauter. 1998. The roles of time of day and sleep quality in modulating glucose regulation: clinical implications. *Hormone Research in Paediatrics* 49, 3-4 (1998), 191–201.

[46] Sandra I Sobel, Peter J Chomentowski, Nisarg Vyas, David Andre, and Frederico GS Toledo. 2014. Accuracy of a novel noninvasive multisensor technology to estimate glucose in diabetic subjects during dynamic conditions. *Journal of Diabetes Science and Technology* 8, 1 (2014), 54–63.

[47] Karine Spiegel, Kristen Knutson, Rachel Leproult, Esra Tasali, and Eve Van Cauter. 2005. Sleep loss: a novel risk factor for insulin resistance and Type 2 diabetes. *Journal of applied physiology* 99, 5 (2005), 2008–2019.

[48] Santhisagar Vaddiraju, Diane J Burgess, Ioannis Tomazos, Faquir C Jain, and Fotios Papadimitrakopoulos. 2010. Technologies for continuous glucose monitoring: current problems and future promises. *Journal of Diabetes Science and Technology* 4, 6 (2010), 1540–1562.

[49] Wikipedia. 2017. Blood sugar. (2017). https://en.wikipedia.org/wiki/Blood_sugar.

Received February 2017