<u>Original Paper</u>

Assessing Elevated Blood Glucose Levels Through Blood Glucose Evaluation and Monitoring Using Machine Learning and Wearable Photoplethysmography Sensors: Algorithm Development and Validation

Bohan Shi[1,2], PhD; Satvinder Singh Dhaliwal[3,4,5], PhD; Marcus Soo[1], BEng; Cheri Chan[6], BSc; Jocelin Wong[1], BEng; Natalie W C Lam[2], BSc; Entong Zhou[2], GCE; Vivien Paitimusa[2], GCE; Kum Yin Loke[2], BSc; Joel Chin[2], BEng; Mei Tuan Chua[6], MN; Kathy Chiew Suan Liaw[6], BSc; Amos W H Lim[1], MSc; Fadil Fatin Insyirah[6], BSc; Shih-Cheng Yen[7], PhD; Arthur Tay[8], PhD; Seng Bin Ang[9,10], MD

[1]Actxa Pte Ltd, Singapore, Singapore

[2]Activate Interactive Pte Ltd, Singapore, Singapore

[3]Curtin Health Innovation Research Institute, Curtin University, Perth, Australia

[4]Faculty of Health Sciences, Curtin University, Perth, Australia

[5]Duke-NUS Graduate Medical School, National University of Singapore, Singapore, Singapore

[6]KK Women's and Children's Hospital, Singapore, Singapore

[7]Innovation and Design Programme, Faculty of Engineering, National University of Singapore, Singapore, Singapore

[8]Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore

[9]Family Medicine Academic Clinical Program, Duke-NUS Medical School, Singapore, Singapore

[10]Menopause Unit, KK Women's and Children's Hospital, Singapore, Singapore

Corresponding Author:
Bohan Shi, PhD
Actxa Pte Ltd
#13-06A SingPost Center
10 Eunos Road 8
Singapore, 408600
Singapore
Phone: 65 88115658
Email: bohan.sbh@gmail.com

Abstract

Background: Diabetes mellitus is the most challenging and fastest-growing global public health concern. Approximately 10.5% of the global adult population is affected by diabetes, and almost half of them are undiagnosed. The growing at-risk population exacerbates the shortage of health resources, with an estimated 10.6% and 6.2% of adults worldwide having impaired glucose tolerance and impaired fasting glycemia, respectively. All current diabetes screening methods are invasive and opportunistic and must be conducted in a hospital or laboratory by trained professionals. At-risk participants might remain undetected for years and miss the precious time window for early intervention to prevent or delay the onset of diabetes and its complications.

Objective: We aimed to develop an artificial intelligence solution to recognize elevated blood glucose levels (≥7.8 mmol/L) noninvasively and evaluate diabetic risk based on repeated measurements.

Methods: This study was conducted at KK Women's and Children's Hospital in Singapore, and 500 participants were recruited (mean age 38.73, SD 10.61 years; mean BMI 24.4, SD 5.1 kg/m^2). The blood glucose levels for most participants were measured before and after consuming 75 g of sugary drinks using both a conventional glucometer (Accu-Chek Performa) and a wrist-worn wearable. The results obtained from the glucometer were used as ground-truth measurements. We performed extensive feature engineering on photoplethysmography (PPG) sensor data and identified features that were sensitive to glucose changes. These selected features were further analyzed using an explainable artificial intelligence approach to understand their contribution to our predictions.

Results: Multiple machine learning models were trained and assessed with 10-fold cross-validation, using participant demographic data and critical features extracted from PPG measurements as predictors. A support vector machine with a radial basis function kernel had the best detection performance, with an average accuracy of 84.7%, a sensitivity of 81.05%, a specificity of 88.3%, a precision of 87.51%, a geometric mean of 84.54%, and F score of 84.03%.

Conclusions: Our findings suggest that PPG measurements can be used to identify participants with elevated blood glucose measurements and assist in the screening of participants for diabetes risk.

Introduction

Diabetes mellitus (DM) is a chronic and heterogeneous metabolic disorder characterized by the presence of hyperglycemia due to deterioration of insulin secretion, defective insulin action, or both [1,2]. There are 3 main types of DM: type 1 DM (T1DM), type 2 DM (T2DM), and gestational diabetes. T2DM is the most prevalent type of diabetes, affecting over 95% of people with diabetes worldwide [3,4].

The prevalence of DM has been proliferating in recent decades, and it is now the most prominent and fastest-growing global public health challenge [5,6]. Uncontrolled diabetes is associated with an increased risk of complications such as cardiovascular disease, kidney failure, vision loss, nerve damage, and overall mortality [7-9]. On the basis of the latest diabetes prevalence estimate, 10.5% of the global adult population is affected by diabetes, and almost half of them are undiagnosed [10]. The growing at-risk population has further strained scarce health resources. Globally, approximately 10.6% of adults have impaired glucose tolerance (IGT) and 6.2% have impaired fasting glycemia (IFG) [4]. IGT and IFG are reversible transitional conditions between normality and diabetes. These conditions, also known as prediabetes, are characterized by elevated blood glucose levels that are not high enough to be classified as diabetes. However, individuals with IGT or IFG are at increased risk of developing cardiovascular disease, coronary heart disease, stroke, and mortality [11]. One of the challenges with IGT and IFG is that they often do not have any obvious symptoms, which means that they can go undetected and undiagnosed for years. Moreover, a follow-up study conducted in Singapore reported that one-third of these individuals with prediabetes would likely develop T2DM within 8 years without lifestyle changes [12]. A similar study with data from the United Kingdom has also reported that a substantial proportion of individuals with prediabetes could progress to T2DM within 5 years [13]. Therefore, predicting the risk of diabetes in the asymptomatic population is a significant health challenge that must be addressed. Early recognition of prediabetes and undiagnosed T2DM will result in a better health outcome or a more favorable long-term prognosis [14].

Currently, the diagnosis of diabetes and prediabetes is well established. T2DM and prediabetes can be detected using one of four methods: (1) the fasting plasma glucose value, (2) the 2-hour plasma glucose value during a 75 g oral glucose tolerance test, (3) hemoglobin A_{1c}, and (4) a random plasma glucose test [3]. All these diagnostic screening methods are invasive and opportunistic in nature and must be conducted in a hospital or laboratory by trained professionals. A confirmed diagnosis usually requires repeated testing. As all the tests are single-time point screenings, adults aged >35 years are recommended to undergo regular screening every 3 years. Nevertheless, at-risk individuals hardly comply with this recommendation, especially in developing countries, owing to the cost of diagnostic tests and the scarcity of medical resources [15,16].

Unlike T1DM and gestational diabetes, the development of T2DM and its complications is preventable or controllable. A considerable number of studies have shown that lifestyle and behavioral interventions help patients with diabetes achieve adequate glycemic control [17,18]. Recent evidence also suggests that early lifestyle adjustment will help participants with prediabetes return to normoglycemia and reduce the risk of developing T2DM [19-21]. Frequent diabetes screening identifies individuals with a high risk of T2DM 2.2 years earlier [22], creating a precious time frame and opportunity for taking an early intervention to prevent or delay the onset of diabetes and its complications and improve overall clinical outcomes.

For established individuals with diabetes, constant monitoring of their blood glucose concentration is crucial so that appropriate insulin dosage can be administered in a timely manner to avoid acute and chronic complications and delay disease progression. Conventional blood glucose measurement requires patients to prick their fingers several times a day, which causes the development of massive scarring and loss of sensation at the fingertips over the year [23]. This measurement method is invasive, inconvenient, and expensive, which are the main barriers to the effective self-management of diabetes in the older adult group [24,25]. To improve diabetes outcomes and assist patients in self-managing the disease, continuous glucose monitoring devices have entered the market and are made available for some patients with diabetes. However, most continuous glucose monitoring sensors currently available are still invasive, which measures glucose concentration in the subcutis using an electrochemical needle sensor [26]. Users need to replace the sensor frequently and purchase different components of the system regularly, which will cost from US $2500 to US $6000 per year [27,28].

In recent years, the advancement and use of wearable technologies and artificial intelligence (AI) have gradually

changed our daily lives, as many people use wrist-worn wearables daily for fitness and health monitoring [29]. Most consumer wearables have incorporated green light reflection photoplethysmography (PPG) sensors into their products. Wearable technology has the potential to greatly expand the impact of public health initiatives by using a proactive approach to identify abnormal physiological signals, assessing disease risk factors, and helping patients manage chronic conditions and recovery [30-33].

In 2011, Monte-Moreno [34] demonstrated the use of PPG data collected using a pulse oximeter to estimate blood glucose levels. By analyzing the PPG waveform, features such as the respiration frequency, heart rate variability (HRV), and other physiological parameters can be extracted. They are then fed into a random forest model, yielding a prediction accuracy of 87.7% based on the Clark error grid. Rodin et al [35] validated a wearable biosensor developed by Zilberstein et al [36] as an indirect measure of glucometry. The biosensor comprises a PPG sensor and an optically sensitive backglass panel that changes its optochemical characteristics according to the concentrations of specific sweat metabolites. In total, 200 adult participants were recruited, and each participant wore a smartwatch to extract PPG data, while blood samples were collected from the antecubital vein concurrently. The estimation of the blood glucose level was derived using a proprietary algorithm developed by SpectroPhon and compared against a glucose lactate analyzer (YSI 2300). The proposed biosensor was able to detect anteprandial glucose with a mean absolute percentage error of 7.4% and a normalized root mean squared error of 11.56%, while postprandial glucose measurements yielded 7.54% mean absolute percentage error and 9.79% normalized root mean squared error. Zhang et al [37] used a smartphone, taking a video of the index finger covering the flash, to capture the fluctuation in the light absorption associated with the change in blood volume. The resulting red, green, and blue image was then transformed into PPG data. The Gaussian fitting method was applied to model the PPG waveform components, from which 28 time-domain and frequency-domain features were extracted. A support vector machine (SVM) with a Gaussian kernel was trained with data from 80 participants to classify the user's glucose level as normal, borderline, or warning, with an accuracy of 81.49%, 79.85% sensitivity, 83.19% specificity, and 80.2% F score. The study was conducted in a highly controlled environment with limited participants, so the generalizability of these results is subject to certain limitations.

Conventional blood glucose monitoring technologies often require invasive measures such as finger pricking or the use of skin sensors and patches. These methods can be uncomfortable and inconvenient for users and can also be financially burdensome. To address these issues, we propose a novel solution called blood glucose evaluation and monitoring (BGEM) that leverages the latest advancements in signal processing, wearable technology, and AI to detect elevated blood glucose levels and evaluate the risk of developing diabetes. With BGEM, users only need to measure their PPG data using a consumer-grade wrist-worn wearable device. The AI model will then compute relevant digital biomarkers and evaluate the risk of prediabetes or T2DM by recognizing elevated blood glucose levels (≥7.8 mmol/L). This solution allows for frequent blood glucose testing without the discomfort and inconvenience of current technologies.

Methods

PPG Sensor

PPG is a low-cost, noninvasive technique that measures the volumetric fluctuation in arterial blood flow [38]. The human wrist is one of the sites for measuring the PPG signal because it has a rich arterial source and an excellent sensor placement with minimal interference to one's daily activities. The PPG signal comprises superimposed pulsatile alternating current components and direct current voltage components. A PPG signal is obtained by illuminating the light emitting device on the skin surface and measuring the variations in light absorption or reflection that reflect the pulsatile flow patterns, as shown in Figure 1.

Figure 1. Illustration of the working principle of a photoplethysmography (PPG) sensor. Changes in blood flow represent different phases within the cardiac cycle. During the diastolic phase, blood volume, arterial diameter, and hemoglobin concentration in the measurement site are minimized, leading to minimum absorption of light by blood and, consequently, an increase in light intensity detected by the sensor system. The reverse is valid for the systolic phase, where a decrease in light intensity is detected instead. AC: alternating current; DC: direct current.
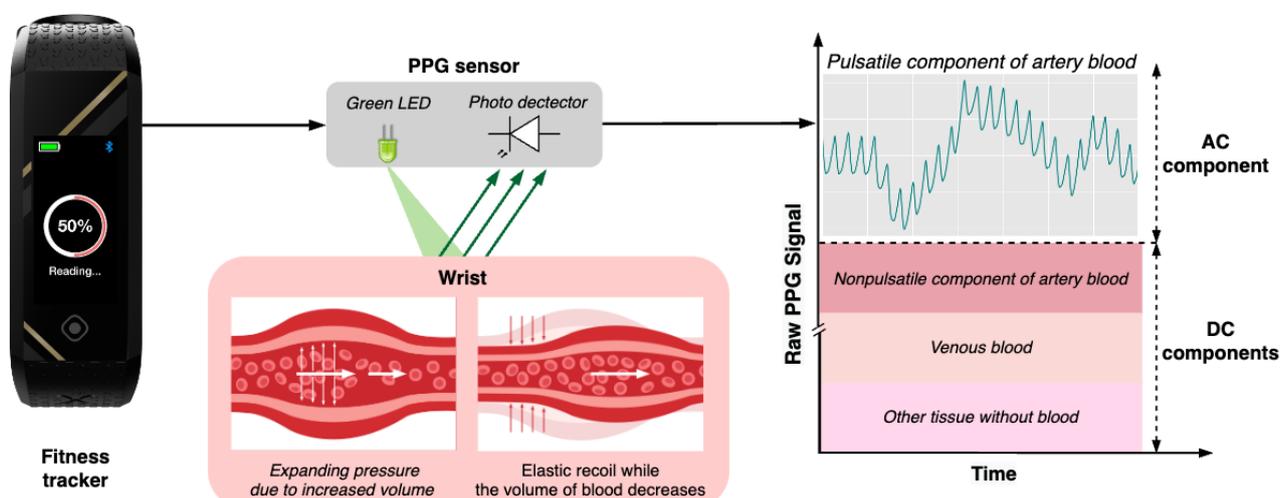
The pulsatile alternating current component corresponds to the cardiac cycle, characterizing that the wrist's blood vessels expand and contract with each heartbeat, whereas the direct current component reflects constant light absorption by venous and arterial blood, as well as other tissues [39]. The PPG signal can detect vascular changes associated with diabetes and contains substantial valuable information from HRV, which is significantly associated with diabetes [40]. Hence, it will be used in this study to extract valuable and meaningful features to identify an individual's glucose status (elevated or normal).

Ethical Considerations

Before commencing the study, ethical clearance was obtained from the SingHealth Centralised Institutional Review Board of Singapore (2020/2968) on March 21, 2021. All methods were performed in accordance with Singapore's clinical guidelines and regulations. Informed consent was obtained from all the trial participants or their legal guardians. The clinical trial was registered on ClinicalTrials.gov (NCT05504096) on August 17, 2022.

Study Protocol

In total, 500 participants were recruited from KK Women's and Children's Hospital in Singapore. Participants' demographics are summarized in Table 1. For most participants, the blood glucose levels were measured before and after consumption of 75 g of a sugary drink using both the conventional glucometer (Accu-Chek Performa) and the wrist-worn wearable device. Participants who were excluded for the second measurement had high blood glucose measurements ≥11.1 mmol/L on their first measurement and hence were not administered the sugary drink measuring 75 g.

After consuming the sugary drink, 55.1% (266/483) of the participants had high blood glucose (≥7.8 mmol/L). The distribution of blood glucose levels before and after consuming the sugary drink is shown in Figure 2. A statistically significant difference was observed between the 2 distributions ($P<.001$).

Table 1. Description of participants (N=500).

Characteristics	Values
Demographic data	
Age (years), mean (SD); range	38.73 (10.61); 21-81
BMI (kg/m^2), mean (SD); range	24.4 (5.1); 16.3-71.1
Gender, n (%)	
Men	51 (10.2)
Women	449 (89.8)
Diabetes profile	
Family history of diabetes, n (%)	
Yes	157 (31.4)
No	343 (68.6)
Prediabetes, n (%)	
Yes	17 (3.4)
No	483 (96.6)
Diabetes, n (%)	
Yes	8 (1.6)
No	492 (98.4)
Gestational diabetes, n (%)	
Yes	21 (4.2)
No	428 (85.6)
N/A[a]	51 (10.2)

[a]N/A: not applicable.

Figure 2. The distribution of ground-truth blood glucose levels before and after sugary drinks (*P*<.001).
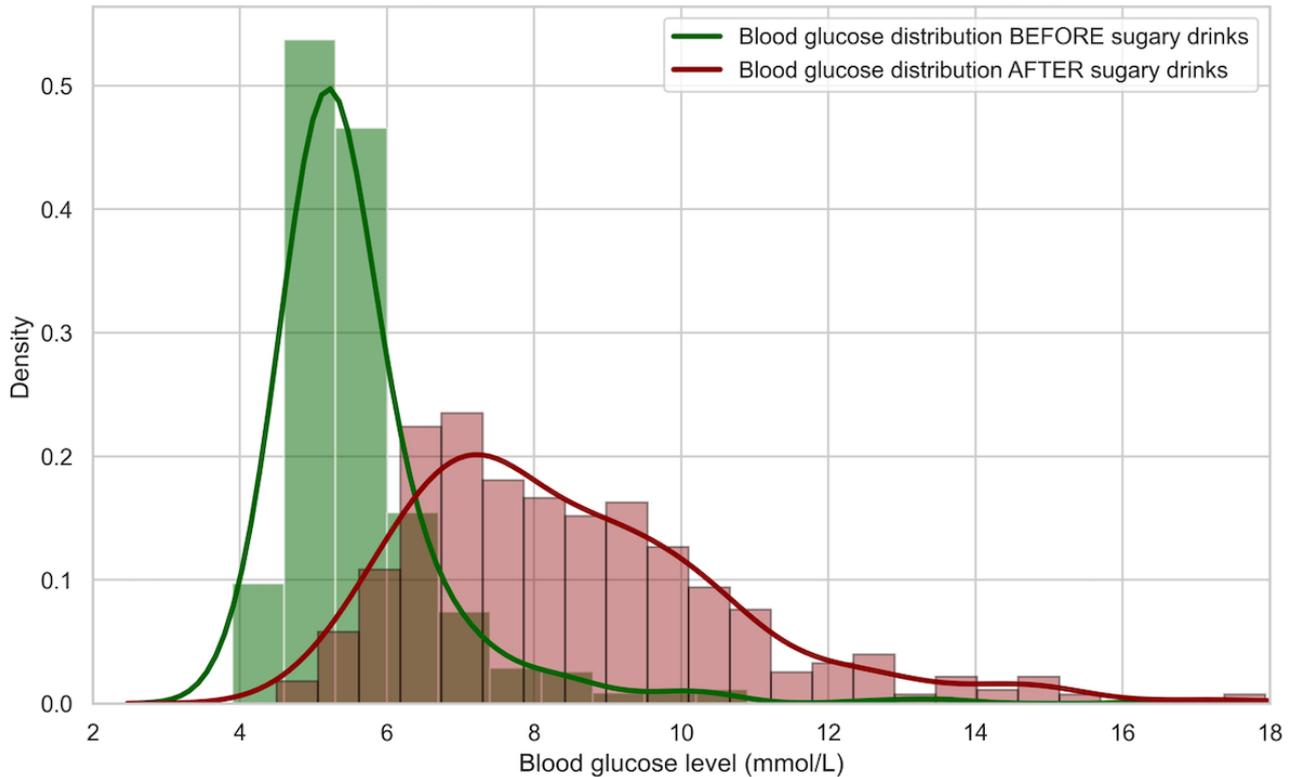
Study Device

The Actxa Spark+ Series 2, a low-cost and commercially available wrist-worn wearable device, was used in this project. This multifunctional device, built for everyday activities, fitness, and preventive health monitoring, provided an adequate PPG signal quality at 50 Hz. The wearable device is equipped with advanced PPG technology that enables accurate and reliable measurement of heart rate (HR) and other physiological parameters. This is similar to the devices used in Singapore's nationwide health care campaigns, such as the National Steps Challenge. It is also worth noting that our proposed solution is device agnostic and can be easily integrated into other wearables with PPG capabilities, allowing for a scalable and cost-effective assessment of risk-based populations, including high-risk participants, participants with undiagnosed diabetes, and patients in need of primary prevention interventions.

Before Processing

The raw PPG signal was collected using both wrist-worn wearables in 16-bit binary format. We first performed a digital-to-analog conversion using the following formula:

$$V_i = 5 \times \frac{Signal_i}{2^{16}} \quad (1)$$

Liang et al [41] suggested that a fourth-order Chebyshev II filter provides an optimal processing performance for short PPG signals. Hence, we adopted the recommended filter design to remove low-frequency drift and high-frequency noise using a band-pass Chebyshev II filter. The proposed band-pass filter has a lower cut-off frequency of 0.3 Hz and an upper cut-off frequency of 4 Hz.

The filtered PPG signals still contain various forms of outliers, such as peaks with abnormally high amplitudes or distortions in the oscillating waveform, which can be caused by movement from the upper extremity or improper contact between the sensor and skin. Features derived from signals that possess outliers may not be accurate, so a *z* scores outlier detection with a cut-off value of 3 SDs of the mean was applied. The identified outliers or regions of outliers were replaced with a reasonable estimate via a nearest neighbor interpolation for the HRV feature extraction. Because PPG signals do not change drastically in such a short duration, this method is determined to be an appropriate approach to the problem. Furthermore, the number of outliers was minimal in our data set, and hence should not have affected the features that we generated later. The data preprocessing steps are illustrated in Figure 3.

Figure 3. Data preprocessing workflow. (A) Raw photoplethysmography (PPG) signal, (B) removal of the signal's moving trend using a Chebyshev high-pass filter, (C) use of a Chebyshev low-pass filter to eliminate high-frequency noise, and (D) final step involves outlier identification from the filtered PPG signal. DAC: digital-to-analog conversion.
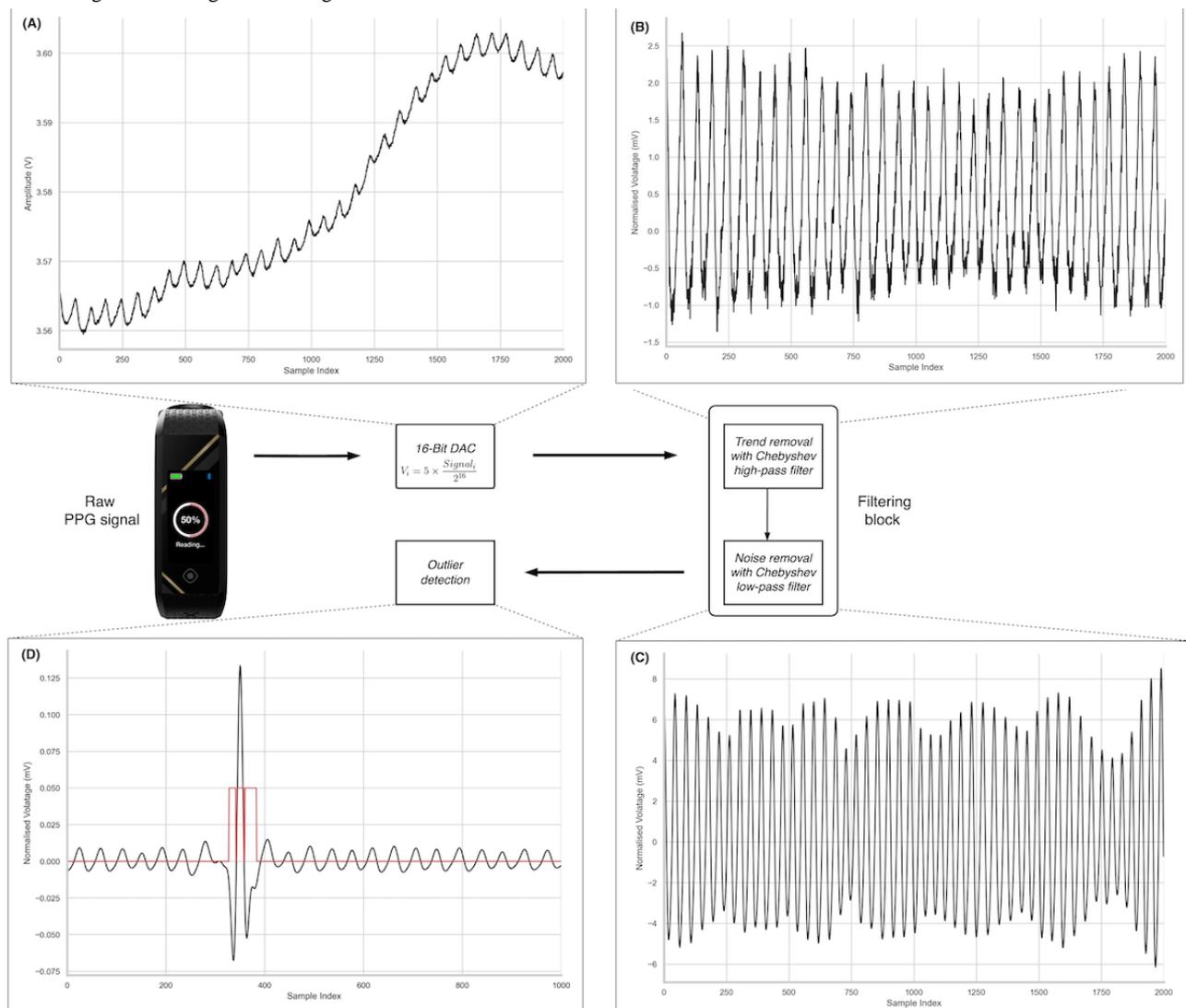
Feature Extraction

Overview

The preprocessed data were suitable for generating reliable features, and a total of 248 features were generated. These features can be classified into seven categories: (1) HRV features, which encompass time domain, frequency domain, and nonlinear HRV features; (2) waveform features; (3) HR features; (4) energy measure features; (5) complexity measure features; (6) continuous wavelet transform (CWT) features; and (7) patient demographics. The complete set of features analyzed in this study is summarized in Multimedia Appendix 1. However, these 248 feature candidates are not all relevant to the change in glucose level, and redundant features might cause prediction performance deterioration. The details of the feature-engineering and feature-selection process are discussed in the "Feature Selection" section.

HRV Features

HRV is the variation in time intervals between consecutive heartbeats and is widely used as a noninvasive physiological biomarker of the autonomic nervous system response [42-44]. HRV provides a proxy to measure sympathetic nervous system (SNS) and parasympathetic nervous system (PNS) activity, which reflects the ability to respond to and recover from abrupt physical, psychological, and environmental changes [44-46]. As HR estimated at any given time represents the net effect of the neural output of the PNS, which slows HR, and SNS, which accelerates HR, HRV also detects imbalance in the autonomic nervous system resulting from over- or understimulation of SNS and PNS. Therefore, the fluctuation in HRV values provide useful insights into many clinical applications, such as mental stress, exercise and rehabilitation, cardiovascular fitness, pathological state, progression of chronic disease, and even predicting the onset of diseases [47-51]. Depending on the application, HRV features are usually extracted from an ultra–short-term (<5 min), short-term (approximately 5 min), or whole-day 24-hour time frame [52]. Most HRV features can be grouped under time-domain, frequency-domain, or nonlinear categories. In this project, most of the widely used HRV features were included in our analysis and were extracted using a

5-minute time frame. These HRV features are briefly explained in Multimedia Appendix 1 using the feature indices (F1-F71).

HR Features

Prior studies have noted the influence of impaired blood glucose on HR, especially resting HR [53,54]. Hence, HR was extracted by finding the number of peaks for every 10 seconds of the filtered PPG signal. The statistical features of the HR were then calculated and used as part of the feature inputs (F72-F81).

Wavelet Analysis

A considerable number of studies have applied wavelet transformation to analyze HRV data associated with a wide variety of health care applications. Earlier research has used features derived from CWT to predict blood glucose levels [55]. In this project, we applied CWT to the PPG signal using the Mexican Hat mother wavelet. The mean, SD, and maximum value of the resulting CWT matrix were included in the feature vector (F82-F84).

Waveform Features

Previous studies have reported that the characteristics of the PPG waveform extracted from healthy participants and participants with diabetes exhibited statistical differences [37,56]. Nirala et al [56] also suggested that the first and second eigenvalues derived from the first derivative of the PPG signal are the top features for identifying T2DM. In addition, several studies have revealed a functional relationship between the PPG signal and blood glucose levels [34,57]. Similarly, respiratory information can also be extracted from the PPG waveform [33,58]. However, PPG waveforms derived from signals using a wrist-worn PPG sensor often have a nondetectable diastolic peak and a dicrotic notch, unlike the signals collected using fingertip PPG.

Waveform features (F85-F196) derived from the PPG waveform were included in the feature set, and the definition of the waveform features is illustrated in Figure 4.

Figure 4. Definition of the photoplethysmography (PPG) waveform features. AUF: area under the falling edge; Apulse: area under a PPG wave; AUR: area under the rising edge; FN: magnitude of falling edge; Fslope: slope of falling edge; FT: fall time; RP: magnitude of rising edge; Rslope: slope of rising edge; RT: rise time.
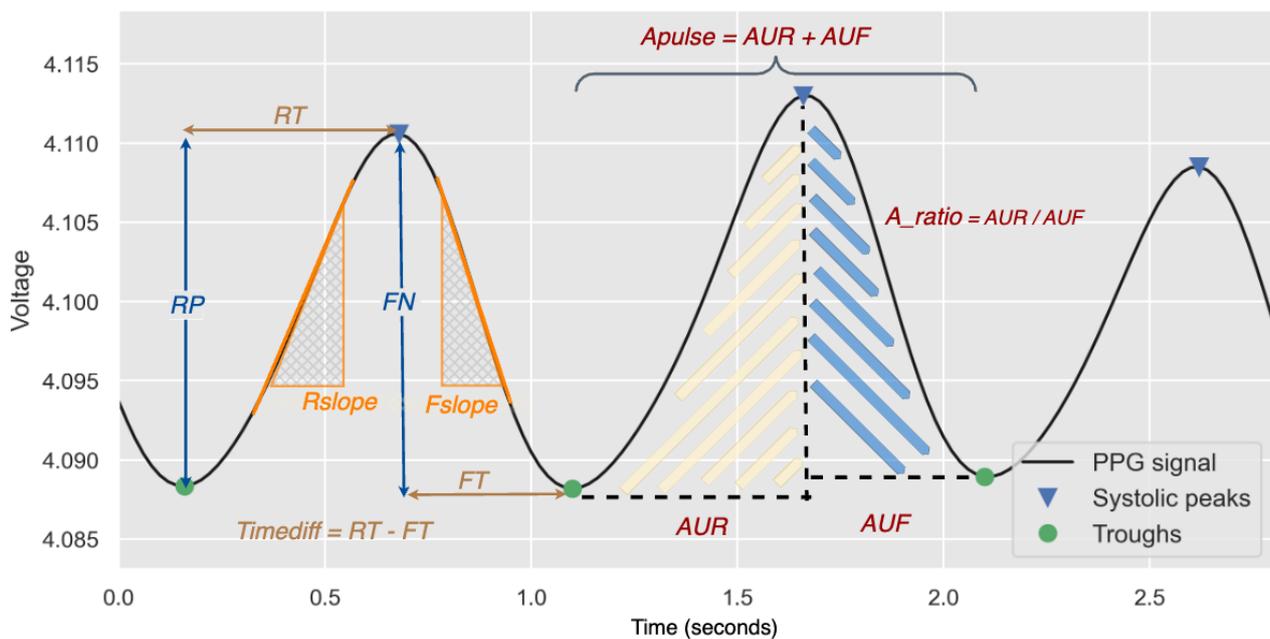
Energy Measures

Several studies have used the energy features extracted from PPG signals to estimate blood glucose [34,59,60]. The Kaiser-Teager energy (KTE) operator and logarithmic energy are 2 commonly used methods to analyze the energy profile. These features were computed from a 5-second sliding window, as it ensures that the PPG signals within each window would be long enough to contain several heartbeats but short enough such that the wave amplitude changes are negligible.

The KTE operator is a well-known method for providing a time-frequency analysis of the instantaneous energy of the PPG signal from the amplitude and frequency. Using the implementation strategy explained by Monte-Moreno [34], we computed the energy profile of the PPG signal at each sliding

window frame, and the KTE operator for the n-th frame was computed using the following equation:

$$KTE_n(i) = x_{frame}(i)^2 - x_{frame}(i+1) * x_{frame}(i-1), \text{ which holds for } i = 2,3,...,(L_{frame} - 1) \quad (2)$$

Where x_{frame} is the filtered PPG signal within each sliding window frame.

The statistical metrics were computed for each frame, and the average of the metrics for the nth frame was then calculated and represented as F197 to F206.

To estimate the respiration rate from the PPG signal, we used the logarithmic energy value calculated at the frame level using the following equation:

$$LogE_n = log_2(\sum_{\tau=1}^{L_{frame}} x_{frame}(\tau, n)^2) \quad (3)$$

Where x_{frame} is the filtered PPG signal within each sliding window frame.

The autoregressive model coefficients of order 7 were estimated using the Yule-Walker method, and the Python function *aryule* was used for this purpose. In addition, other statistical parameters were also computed (F207-F223).

Complexity Measures

Sample entropy (SampEn, F224) measures the unpredictability of physiological signals and is commonly used in HRV analysis [61]. The lower the SampEn, the more regular the signal.

SampEn can be defined after calculating the template vector φ^m that is the probability that 2 sequences will match for m points without allowing self-counting [62]:

$$SampEn(m, r, N) = -ln\,[\phi^m(r) - \phi^{m+1}(r)], \quad \textbf{where} \quad \phi^m(r) = \frac{1}{N-m-1}\sum_{i=1}^{N-m+1} C_i^m(r) \quad (4)$$

Where *m* denotes the embedding dimension, tolerance r equals $0.1 * SD$, N denotes the number of data points, and C_i^m counts, within the tolerance resolution r, the number of matching blocks across different embedding dimensions.

SampEn is a tool used to analyze physiological time-series data, but it does not evaluate the complexity of the data at different time scales. Hence, we applied multiscale entropy (MSE) analysis on raw PPG signals to evaluate the hypothetical difference in signal complexity across various time scales for normoglycemia and elevated glucose levels. However, the scale factor was inversely proportionate to the number of data points. From our empirical results, we found that a minimum of 240 pulse waves were required to correctly compute the MSE values over all the timescale factors (τ=20). We found that the sample entropy calculated from PPG signals during periods of elevated blood glucose was significantly higher than that of blood glucose in the normal range at timescale factors between 8 and 14 (τ). This information was then used to create features for the detection of elevated blood glucose levels. Each timescale factor between 8 and 14 was used as a separate feature. In addition, the mean of the adjacent timescale factors was derived to create additional features. These MSE features are represented in the feature vector with feature indices F225 to F244.

Results

Software

All experiments and analyses were performed using Python (version 3.9) and relevant libraries (Table 2). The final model was deployed on Amazon Web Services.

Table 2. A list of the software, and relevant libraries, along with the versions used.

Library	Version
Python	3.9.10
Imbalanced-learn	0.10.1
Joblib	1.2.0
Jupyter	1.0.0
Lightgbm	3.3.4
Matplotlib	3.6.2
Neurokit2	0.2.2
Nolds	0.5.2
Numba	0.53.1
Numpy	1.23.5
Pandas	1.5.2
Pillow	9.3.0
Scikit-learn	1.1.3
Scipy	1.8.0
Seaborn	0.12.1
Spectrum	0.8.1
Statsmodels	0.13.5
Xgboost	1.7.2

Feature Selection

Considering AI ethics and the practicality of implementing the algorithm, some demographic data, such as skin color, race, and personal lifestyle habits, were not used as inputs to the models. However, other general personal characteristics associated with the risk of developing T2DM, such as age,

XSL•FO

RenderX

gender, BMI, and family health history of diabetes, were added to the feature vector before the feature-selection process.

The redundant or irrelevant features might hinder the performance of the prediction model. To reduce the dimensionality of the input features, we applied an ensemble strategy that uses multiple feature-selection algorithms. This creates an optimal feature subset that minimizes the prediction error rate and is the most relevant for predicting the target variable. The ensemble feature-selection steps are summarized as follows:

- Six feature-selection methods, including ANOVA correlation coefficient, mutual information, dispersion ratio, recursive feature elimination, lasso regression, and Extreme Gradient Boosting, were used to choose the 30 best features independently.
- We combined the features obtained from each feature-selection method and ranked them using a majority vote approach to find the common features selected by more than 1 model.
- The highly correlated features were dropped from the selected feature subset.

In total, 12 features were selected from the entire feature set and ranked based on the results of the feature-selection strategy (Table 3). In our study, these selected features were the most sensitive predictors for capturing the characteristics of a participant's elevated blood glucose levels.

Table 3. The selected top features after the ensemble feature-selection method.

Rank	Feature
1	Welch_hf_rel
2	AR_hf_rel
3	A_FE_mean
4	A_ratio_mean
5	Age
6	A_Pulse_iqr
7	KTE_skew
8	LOG_std
9	BMI
10	MSE_sum_13_14
11	Family history
12	A_ratio_max
13	Gender[a]

[a]Note that gender was not selected as a top feature in our feature-selection algorithm. However, it was previously identified as a sensitive predictor for T2DM, in which the prevalence of T2DM in men was higher than that in women [63]. This discrepancy could be attributed to the gender imbalance in the data set (men: 10.2%; women: 89.8%). Therefore, we included gender as one of the top features to provide a complete user profile for future investigation and development.

The selected features could be further divided into 4 main categories. Under the time-domain features, the selected features were the area under the PPG curves. A_FE_mean refers to the average area under the falling edge of each pulse (Figure 4). A_ratio refers to the ratio of the area under the rising edge to the area under the falling edge of each pulse (Figure 4), and both the average and maximum values were deemed relevant to the model's predictions. A_pulse_iqr refers to the IQR of the total area under each pulse (Figure 4). In the frequency domain, the selected features were the relative powers of the high-frequency bands in both the Welch power spectral density (PSD; Multimedia Appendix 1, F32-F44) and autoregressive PSD (Multimedia Appendix 1, F45-F57).

In the nonlinear domain, the selected features were either related to the energy or the complexity of the signal. LOG_std refers to the SD of log-energy entropy (equation 3), whereas KTE_skew refers to the skewness of the KTE energy measure for each sliding window (equation 2). Furthermore, the complexity feature that was selected was the sum of the MSE over 2 scales, 13 and 14.

Finally, the remaining selected features were demographic features that described the age and BMI of the participants, as well as if they had any family history of diabetes.

Machine Learning Model Performance

Seven widely used machine learning (ML) algorithms, including the naive Bayes classifier, K-nearest neighbors algorithm, logistic regression, random forest, SVM, XGB, and light gradient boosting machine, were trained with the selected features as inputs. We fine-tuned the hyperparameters of each model and validated their performance using the stratified 10-fold cross-validation method. We adopted multiple regularization techniques across various models to prevent overfitting during the model training. Six evaluation metrics, accuracy, sensitivity, specificity, precision, geometric mean (G-mean), and *F* score, were used to evaluate the model's performance, as accuracy alone cannot provide a comprehensive

examination of model performance due to data imbalance. The G-mean and *F* score are critical evaluation criteria to assess the models' performance, as they are robust to significant label imbalance.

The prediction results from each model are reported as the mean and SD of the evaluation metrics, and Table 4 shows the

summary of the results. SVM with the radial basis function kernel showed the best prediction performance with an average accuracy of 84.7%, a sensitivity of 81.05%, a specificity of 88.35%, and a precision of 87.51%. In particular, the average G-mean was 84.54% and *F* score was 84.03%.

Table 4. The prediction results obtained from 10-fold cross-validation using various machine learning models.

Model	Accuracy		Sensitivity		Specificity		Precision		Geometric mean		*F* score	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
NB[a]	60.51	4.63	66.17	7.44	54.87	5.78	59.43	4.12	60.08	4.6	62.51	5.19
KNN[b]	76.7	3	90.45	4.30	62.94	4.15	70.97	2.47	75.4	3.09	79.5	2.68
LR[c]	63.1	4.65	64.56	7.07	61.66	4.30	62.65	4.16	63	4.67	63.52	5.37
RF[d]	76.76	5.73	76.84	8.18	76.69	6.42	76.81	6.08	76.64	5.72	76.68	6.23
SVM[e]	84.7	4.14	81.05	6.77	88.34	4.19	87.51	4.26	84.54	4.18	84.03	4.58
XGB[f]	78.06	4.91	77	6.58	79.12	4.98	78.7	4.88	78	4.89	77.77	5.15
LGBM[g]	77.9	3.98	75.54	7.36	80.27	4.45	79.35	4.1	77.74	4.07	77.24	4.81

[a]NB: naive Bayes.

[b]KNN: K-nearest neighbors.

[c]LR: logistic regression.

[d]RF: random forest.

[e]SVM: support vector machine.

[f]XGB: Extreme Gradient Boosting.

[g]LGBM: light gradient boosting machine.

Model Interpretation

The use of deep learning in the medical and health care domain has shown great potential for solving a range of problems, such as detecting specific symptoms or abnormalities [64,65]. However, the interpretability of deep learning models remains a significant challenge, and it is often difficult for clinicians to trust the decisions made using a black-box system. The lack of model interpretability also raises ethical concerns, particularly when the decision fails. Furthermore, our current data set is considerably small (500 participants) compared with typical deep learning models in other domains, which are trained with thousands of data points. Deep learning models are known to perform well with a larger data set and fail to learn meaningful representations when there is a lack of data [66]. Therefore, we did not investigate the use of deep learning in this study.

As the proposed ML model is designed to complement the existing diabetes detection solution and is relatively new to the clinical community, the features selected in the previous section must be interpretable and exhibit a certain level of agreement with existing findings. A family history of diabetes, being male, being aged ≥45 years, and having an increased BMI have been

identified as major risk factors in the literature for developing prediabetes or T2DM [63,67,68]. These 4 risk factors were part of the selected predictors, and this paper provides a preliminary attempt to explain how the selected predictors contribute to detecting elevated blood glucose using the Shapley additive explanations (SHAP) framework. SHAP is a game theoretical approach that provides global and local explanations of the association between the ML output and input features [69].

Figure 5A illustrates the SHAP values of each feature across all the predictions from the training set. The features were ranked by their mean SHAP values, with larger values shown in red and smaller values shown in blue. The beeswarm plot revealed that a family history of diabetes, increasing age, and higher BMI are associated with a higher probability of elevated blood glucose levels. These observations are consistent with previous research and demonstrate that the ML algorithm has successfully captured the relationship between these features and elevated blood glucose levels. In addition, other proposed features showed varying levels of impact on the model's output. However, the gender feature did not have any apparent effect on the model's predictions.

XSL•FO
RenderX

Figure 5. The Shapley additive explanations (SHAP) plots indicate the association between the selected features and their impact on the predicted outcome. (A) SHAP beeswarm plot and (B) SHAP waterfall plot.
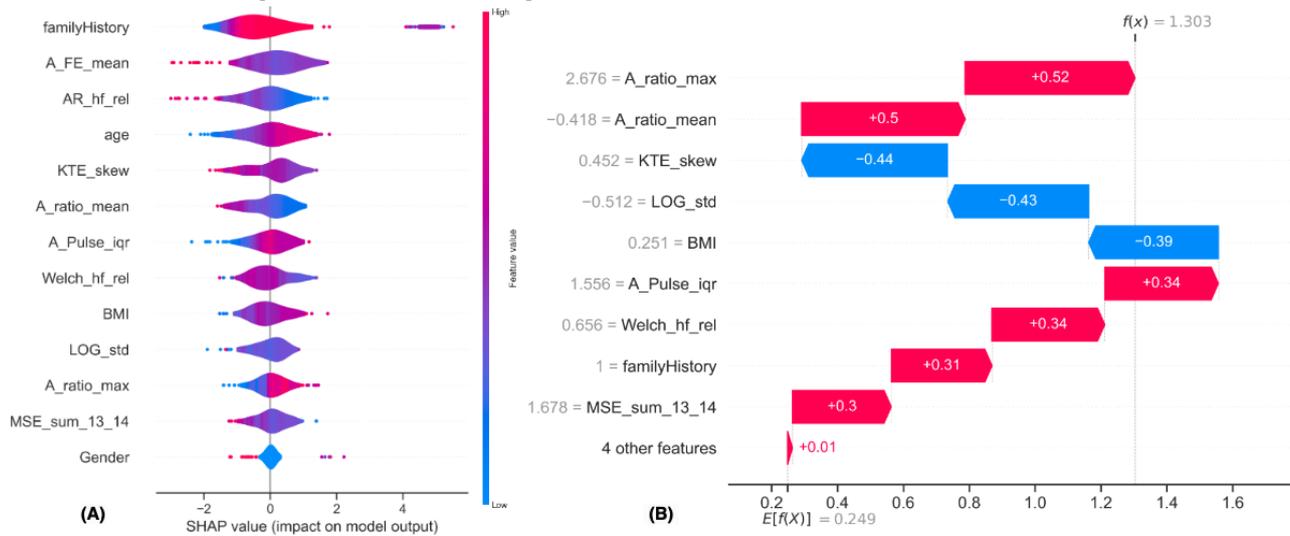
In Figure 5B, each row in the plot shows how the contributions of different features move the output of the model from the expected value ($E[f(x)]$) to the actual prediction output $f(x)$ for a single sample with a positive class prediction (blood glucose level ≥7.8 mmol/L) in the test set. The expected value, $E[f(x)]$, is determined using the entire training data set. As expected, most features provide positive SHAP values in this sample, which collectively push the model's output toward the correct prediction. However, this specific test participant's BMI was in the healthy range, which pushed the model's output toward the normal class and might have resulted in a false negative prediction. This indicates that relying on a single feature or demographic data alone may not provide an accurate prediction of blood glucose levels.

Using the SHAP values, we can understand the model's overall behaviors and how features affect the output positively or negatively, which can help improve the prediction model in the future.

Assessment of the Elevated Blood Glucose Levels From Multiple Measurements

Generally, diagnostic tests are not highly sensitive and highly specific. Therefore, repeated measurements of the wrist-worn wearable device were combined and assessed in an optimum fashion to maximize sensitivity, specificity, and precision.

Consecutive measures of blood glucose were combined in parallel using the "AND" and "OR" rules to assist in the detection of elevated blood glucose measurement levels. The "OR" rule increases the overall sensitivity, and the "AND" rule increases the overall specificity, which is greater than that of either test alone [70].

Discussion

Principal Findings

While the health care landscape is changing, the rapidly aging society and the need for improved population health outcomes call for new models of care to effectively prevent the onset and delay the progression of chronic diseases. Furthermore,

short-term health behaviors contribute significantly toward long-term health outcomes, while unattended and frequent glucose spikes might result in prediabetes and eventually diabetes. The availability of noninvasive and device-agnostic blood glucose detection solutions will allow for more frequent and better monitoring of blood glucose levels, thereby reducing the risk of developing T2DM. This study demonstrates that a noninvasive method of assessing diabetes risk using PPG is a viable option to provide a cheaper and accessible modality for the population-wide screening of blood glucose levels. This population-based screening would allow for the earlier detection of DM in the population, especially among those individuals who are unaware of their elevated blood glucose levels. Hence, timely and appropriate lifestyle advice and medical interventions can be provided to prevent diabetes complications. This will subsequently reduce the health care burden for both the individual and the society.

BGEM is a cloud-based solution that can frequently monitor multiple digital biomarkers with minimal disruption to daily life. Developed using the advanced ML operations practice, BGEM can be easily scaled to meet the increasing demand for health care services. The solution includes a user-friendly mobile app that can screen a large population to identify high-risk individuals, people with undiagnosed diabetes, and those who require primary prevention intervention. It also provides timely feedback to users through the app, informing them of their diabetes risk and providing targeted, actionable insights to empower them to take a proactive approach to monitor their glucose levels.

Limitations

Our pilot study has certain limitations. Since fasting blood glucose measurements were excluded and the criteria to define normal and abnormal levels under fasting conditions differed from our current cut-off, we must refrain from definitively concluding that our model is applicable to fasting conditions. Regarding gender, our feature-selection model did not specifically incorporate it, and our analysis using SHAP demonstrated that gender exerted minimal influence on model predictions. Moreover, all analyses were adjusted for the

covariate gender, as required. Therefore, we considered gender to have a limited impact and is not a primary limitation of our findings. To address these limitations, we are actively planning the subsequent phase of data collection. This phase will involve collecting fasting blood glucose measurements in a primary care setting, also allowing for a more balanced gender distribution. More importantly, we could expand our participant pool to encompass participants with prediabetes and diabetes. By addressing these gaps, we aimed to offer a more comprehensive and robust assessment of our model's applicability and effectiveness.

There was no longitudinal follow-up of the participants. External validation of our model on an independent sample must be undertaken to further assess the detection accuracy and generalizability of the results. Nevertheless, as a preliminary investigation, the potential implications of our findings are significant as they might offer a means to identify previously undiagnosed prediabetes or diabetes cases at the population level. We anticipate that our study will serve as a foundational stepping stone, paving the way for more comprehensive diabetes research using AI and wearable devices. To the best of our knowledge, there is no publicly available data set that systematically examines the relationship between PPG data and blood glucose levels. Acquiring a substantial volume of data is imperative, encompassing a diverse and representative sample spanning the entire spectrum of glucose values and incorporating relevant sociodemographic factors. Such comprehensive data can be obtained through a collaborative effort involving research institutions and industry partners while ensuring strict adherence to local ethical considerations and data privacy regulations.

We demonstrated that the cloud-based ML model can detect elevated blood glucose levels, where consecutive measurements can be combined in an optimal manner to provide high sensitivity, specificity, and precision. However, further research is required to address these limitations.

Conclusions

In this study, we performed sophisticated feature engineering and found that the features derived from the MSE analysis of PPG signals effectively detect blood glucose changes. We will discuss this set of novel features in detail in a separate paper. To reduce bias and evaluate the generalizability of the model, we used a 10-fold cross-validation to assess its performance. The SVM with the radial basis function model performed the best, with an average accuracy of 84.7%, a G-mean of 84.54%, and an F score of 84.03%. Previous models were developed using smaller samples and have lower model performance measures. Our model was developed with a larger sample of 500 participants, and most participants were assessed before and after the consumption of a sugary drink. It also achieved better detection accuracy.

Authors' Contributions

BS contributed to the study design, conducted the data analysis and experiments, developed the algorithms and models, and drafted the manuscript. SSD designed the study, performed the statistical data analysis, and drafted the manuscript. JW was responsible for model deployment and developed the data pipeline infrastructure. CC, MTC, KCSL, and FFI contributed to data collection. NWCL, EZ, KYL, and VP assisted with the development of the algorithms and supported data collection. AWHL performed data analysis and edited the manuscript. MS contributed to the study design and supervised the study. JC, S-CY, and AT supervised the study. SBA contributed to study design and supervised the study. All authors have reviewed the manuscript.

Conflicts of Interest

The authors would like to disclose that BS, MS, JW, AWHL, and JC are employed by Actxa Pte Ltd. The authors have an approved plan for managing any potential conflicts arising from employment. SBA and SSD are on the advisory board of Actxa Pte Ltd. All other authors declare that they have no conflicts of interest.

Multimedia Appendix 1

Features summary.
[DOCX File , 21 KB-Multimedia Appendix 1]

References

1. National Diabetes Data Group. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. Diabetes 1979 Dec 1;28(12):1039-1057 [doi: 10.2337/diab.28.12.1039] [Medline: 510803]
2. Kerner W, Brückel J. Definition, classification and diagnosis of diabetes mellitus. Exp Clin Endocrinol Diabetes 2014 Jul 11;122(7):384-386 [doi: 10.1055/s-0034-1366278] [Medline: 25014088]
3. American Diabetes Association. 2. Classification and diagnosis of diabetes: standards of medical care in diabetes—2020. Dia Care 2019 Dec 20;43(Supplement 1):S14-S31 [FREE Full text] [doi: 10.2337/dc20-s002]
4. IDF diabetes atlas 10th edition. International Diabetes Federation. 2021. URL: https://diabetesatlas.org/idfawp/resource-files/2021/07/IDF_Atlas_10th_Edition_2021.pdf [accessed 2023-06-04]

5.   Emerging Risk Factors Collaboration; Sarwar N, Gao P, Seshasai SR, Gobin R, Kaptoge S, et al. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. Lancet 2010 Jun 26;375(9733):2215-2222 [FREE Full text] [doi: 10.1016/S0140-6736(10)60484-9] [Medline: 20609967]

6.   Lin X, Xu Y, Pan X, Xu J, Ding Y, Sun X, et al. Global, regional, and national burden and trend of diabetes in 195 countries and territories: an analysis from 1990 to 2025. Sci Rep 2020 Sep 08;10(1):14790 [FREE Full text] [doi: 10.1038/s41598-020-71908-9] [Medline: 32901098]

7.   Li S, Wang J, Zhang B, Li X, Liu Y. Diabetes mellitus and cause-specific mortality: a population-based study. Diabetes Metab J 2019 Jun;43(3):319-341 [FREE Full text] [doi: 10.4093/dmj.2018.0060] [Medline: 31210036]

8.   Saran R, Li Y, Robinson B, Ayanian J, Balkrishnan R, Bragg-Gresham J, et al. US renal data system 2014 annual data report: epidemiology of kidney disease in the United States. Am J Kidney Dis 2015 Jul;66(1):S1-305 [FREE Full text] [doi: 10.1053/j.ajkd.2015.05.001] [Medline: 26111994]

9.   Lau LH, Lew J, Borschmann K, Thijs V, Ekinci EI. Prevalence of diabetes and its effects on stroke outcomes: a meta-analysis and literature review. J Diabetes Investig 2019 May;10(3):780-792 [FREE Full text] [doi: 10.1111/jdi.12932] [Medline: 30220102]

10.   Sun H, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, et al. IDF diabetes atlas: global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. Diabetes Res Clin Pract 2022 Jan;183:109119 [doi: 10.1016/j.diabres.2021.109119] [Medline: 34879977]

11.   Huang Y, Cai X, Mai W, Li M, Hu Y. Association between prediabetes and risk of cardiovascular disease and all cause mortality: systematic review and meta-analysis. BMJ 2016 Nov 23;355:i5953 [FREE Full text] [doi: 10.1136/bmj.i5953] [Medline: 27881363]

12.   Wong M, Gu K, Heng D, Chew SK, Chew LS, Tai ES. The Singapore impaired glucose tolerance follow-up study: does the ticking clock go backward as well as forward? Diabetes Care 2003 Nov;26(11):3024-3030 [FREE Full text] [doi: 10.2337/diacare.26.11.3024] [Medline: 14578234]

13.   Tabák AG, Herder C, Rathmann W, Brunner EJ, Kivimäki M. Prediabetes: a high-risk state for diabetes development. The Lancet 2012 Jun;379(9833):2279-2290 [FREE Full text] [doi: 10.1016/s0140-6736(12)60283-9]

14.   US Preventive Services Task Force; Davidson K, Barry MJ, Mangione CM, Cabana M, Caughey AB, et al. Screening for prediabetes and type 2 diabetes: US preventive services task force recommendation statement. JAMA 2021 Aug 24;326(8):736-743 [FREE Full text] [doi: 10.1001/jama.2021.12531] [Medline: 34427594]

15.   Manne-Goehler J, Geldsetzer P, Agoudavi K, Andall-Brereton G, Aryal KK, Bicaba BW, et al. Health system performance for people with diabetes in 28 low- and middle-income countries: a cross-sectional study of nationally representative surveys. PLoS Med 2019 Mar 1;16(3):e1002751 [FREE Full text] [doi: 10.1371/journal.pmed.1002751] [Medline: 30822339]

16.   Misra A, Gopalan H, Jayawardena R, Hills AP, Soares M, Reza-Albarrán AA, et al. Diabetes in developing countries. J Diabetes 2019 Mar 12;11(7):522-539 [doi: 10.1111/1753-0407.12913] [Medline: 30864190]

17.   García-Molina L, Lewis-Mikhael A, Riquelme-Gallego B, Cano-Ibáñez N, Oliveras-López MJ, Bueno-Cavanillas A. Improving type 2 diabetes mellitus glycaemic control through lifestyle modification implementing diet intervention: a systematic review and meta-analysis. Eur J Nutr 2020 Jun;59(4):1313-1328 [FREE Full text] [doi: 10.1007/s00394-019-02147-6] [Medline: 31781857]

18.   O'Donoghue G, O'Sullivan C, Corridan I, Daly J, Finn R, Melvin K, et al. Lifestyle interventions to improve glycemic control in adults with type 2 diabetes living in low-and-middle income countries: a systematic review and meta-analysis of Randomized Controlled Trials (RCTs). Int J Environ Res Public Health 2021 Jun 10;18(12):6273 [FREE Full text] [doi: 10.3390/ijerph18126273] [Medline: 34200592]

19.   Tuso P. Prediabetes and lifestyle modification: time to prevent a preventable disease. Perm J 2014;18(3):88-93 [FREE Full text] [doi: 10.7812/TPP/14-002] [Medline: 25102521]

20.   Bansal N. Prediabetes diagnosis and treatment: a review. World J Diabetes 2015 Mar 15;6(2):296-303 [FREE Full text] [doi: 10.4239/wjd.v6.i2.296] [Medline: 25789110]

21.   Magkos F, Hjorth MF, Astrup A. Diet and exercise in the prevention and treatment of type 2 diabetes mellitus. Nat Rev Endocrinol 2020 Oct 20;16(10):545-555 [FREE Full text] [doi: 10.1038/s41574-020-0381-5] [Medline: 32690918]

22.   Simmons RK, Griffin SJ, Lauritzen T, Sandbæk A. Effect of screening for type 2 diabetes on risk of cardiovascular disease and mortality: a controlled trial among 139,075 individuals diagnosed with diabetes in Denmark between 2001 and 2009. Diabetologia 2017 Nov;60(11):2192-2199 [FREE Full text] [doi: 10.1007/s00125-017-4299-y] [Medline: 28831539]

23.   Heinemann L. Finger pricking and pain: a never ending story. J Diabetes Sci Technol 2008 Sep 01;2(5):919-921 [FREE Full text] [doi: 10.1177/193229680800200526] [Medline: 19885279]

24.   Hambling CE, Seidu SI, Davies MJ, Khunti K. Older people with Type 2 diabetes, including those with chronic kidney disease or dementia, are commonly overtreated with sulfonylurea or insulin therapies. Diabet Med 2017 Sep;34(9):1219-1227 [FREE Full text] [doi: 10.1111/dme.13380] [Medline: 28498634]

25.   Mattishent K, Lane K, Salter C, Dhatariya K, May HM, Neupane S, et al. Continuous glucose monitoring in older people with diabetes and memory problems: a mixed-methods feasibility study in the UK. BMJ Open 2019 Nov 18;9(11):e032037 [FREE Full text] [doi: 10.1136/bmjopen-2019-032037] [Medline: 31740472]

26. 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 Sep 22;12(5):1064-1071 [FREE Full text] [doi: 10.1177/1932296818774078] [Medline: 29783897]

27. Funtanilla VD, Candidate P, Caliendo T, Hilas O. Continuous glucose monitoring: a review of available systems. P T 2019 Sep;44(9):550-553 [FREE Full text] [Medline: 31485150]

28. Robertson SL, Shaughnessy AF, Slawson DC. Continuous glucose monitoring in type 2 diabetes is not ready for widespread adoption. Am Fam Physician 2020 Jun 01;101(11):646 [FREE Full text] [Medline: 32463633]

29. Sabry F, Eltaras T, Labda W, Alzoubi K, Malluhi Q. Machine learning for healthcare wearable devices: the big picture. J Healthc Eng 2022 Apr 18;2022:4653923 [FREE Full text] [doi: 10.1155/2022/4653923] [Medline: 35480146]

30. Patel S, Park H, Bonato P, Chan L, Rodgers M. A review of wearable sensors and systems with application in rehabilitation. J Neuroeng Rehabil 2012 Apr 20;9(1):21 [FREE Full text] [doi: 10.1186/1743-0003-9-21] [Medline: 22520559]

31. Rodgers MM, Alon G, Pai VM, Conroy RS. Wearable technologies for active living and rehabilitation: current research challenges and future opportunities. J Rehabil Assist Technol Eng 2019 Apr 26;6:2055668319839607 [FREE Full text] [doi: 10.1177/2055668319839607] [Medline: 31245033]

32. Xie Y, Lu L, Gao F, He SJ, Zhao HJ, Fang Y, et al. Integration of artificial intelligence, blockchain, and wearable technology for chronic disease management: a new paradigm in smart healthcare. Curr Med Sci 2021 Dec;41(6):1123-1133 [FREE Full text] [doi: 10.1007/s11596-021-2485-0] [Medline: 34950987]

33. Iqbal SM, Mahgoub I, Du E, Leavitt MA, Asghar W. Advances in healthcare wearable devices. Npj Flex Electron 2021 Apr 12;5(1):9 [FREE Full text] [doi: 10.1038/s41528-021-00107-x]

34. Monte-Moreno E. Non-invasive estimate of blood glucose and blood pressure from a photoplethysmograph by means of machine learning techniques. Artif Intell Med 2011 Oct;53(2):127-138 [FREE Full text] [doi: 10.1016/j.artmed.2011.05.001] [Medline: 21696930]

35. Rodin D, Kirby M, Sedogin N, Shapiro Y, Pinhasov A, Kreinin A. Comparative accuracy of optical sensor-based wearable system for non-invasive measurement of blood glucose concentration. Clin Biochem 2019 Mar;65:15-20 [FREE Full text] [doi: 10.1016/j.clinbiochem.2018.12.014] [Medline: 30629956]

36. Zilberstein G, Zilberstein R, Maor U, Righetti PG. Noninvasive wearable sensor for indirect glucometry. Electrophoresis 2018 Sep;39(18):2344-2350 [FREE Full text] [doi: 10.1002/elps.201700424] [Medline: 29607521]

37. Zhang G, Mei Z, Zhang Y, Ma X, Lo B, Chen D, et al. A noninvasive blood glucose monitoring system based on smartphone PPG signal processing and machine learning. IEEE Trans Industr Inform 2020 Nov;16(11):7209-7218 [FREE Full text] [doi: 10.1109/tii.2020.2975222]

38. Challoner AV, Ramsay CA. A photoelectric plethysmograph for the measurement of cutaneous blood flow. Phys Med Biol 1974 May;19(3):317-328 [FREE Full text] [doi: 10.1088/0031-9155/19/3/003] [Medline: 4445210]

39. Zhao D, Sun Y, Wan S, Wang F. SFST: a robust framework for heart rate monitoring from photoplethysmography signals during physical activities. Biomed Signal Process Control 2017 Mar;33:316-324 [FREE Full text] [doi: 10.1016/j.bspc.2016.12.005]

40. Schroeder EB, Chambless LE, Liao D, Prineas RJ, Evans GW, Rosamond WD, et al. Diabetes, glucose, insulin, and heart rate variability: the Atherosclerosis Risk in Communities (ARIC) study. Diabetes Care 2005 Mar;28(3):668-674 [doi: 10.2337/diacare.28.3.668] [Medline: 15735206]

41. Liang Y, Elgendi M, Chen Z, Ward R. An optimal filter for short photoplethysmogram signals. Sci Data 2018 May 01;5(1):180076 [FREE Full text] [doi: 10.1038/sdata.2018.76] [Medline: 29714722]

42. van Ravenswaaij-Arts CM, Kollée LA, Hopman JC, Stoelinga GB, van Geijn HP. Heart rate variability. Ann Intern Med 1993 Mar 15;118(6):436-447 [FREE Full text] [doi: 10.7326/0003-4819-118-6-199303150-00008] [Medline: 8439119]

43. Xhyheri B, Manfrini O, Mazzolini M, Pizzi C, Bugiardini R. Heart rate variability today. Prog Cardiovasc Dis 2012 Nov;55(3):321-331 [FREE Full text] [doi: 10.1016/j.pcad.2012.09.001] [Medline: 23217437]

44. Thomas BL, Claassen N, Becker P, Viljoen M. Validity of commonly used heart rate variability markers of autonomic nervous system function. Neuropsychobiology 2019 Feb 5;78(1):14-26 [doi: 10.1159/000495519] [Medline: 30721903]

45. Obrist PA. Cardiovascular Psychophysiology: A Perspective. New York, NY: Springer; 1981.

46. Singh N, Moneghetti KJ, Christle JW, Hadley D, Plews D, Froelicher V. Heart rate variability: an old metric with new meaning in the era of using mHealth technologies for health and exercise training guidance. Part one: physiology and methods. Arrhythm Electrophysiol Rev 2018 Aug;7(3):193-198 [FREE Full text] [doi: 10.15420/aer.2018.27.2] [Medline: 30416733]

47. Prinsloo GE, Rauch HL, Derman WE. A brief review and clinical application of heart rate variability biofeedback in sports, exercise, and rehabilitation medicine. Phys Sportsmed 2014 May 13;42(2):88-99 [FREE Full text] [doi: 10.3810/psm.2014.05.2061] [Medline: 24875976]

48. Billman GE, Huikuri HV, Sacha J, Trimmel K. An introduction to heart rate variability: methodological considerations and clinical applications. Front Physiol 2015 Feb 25;6:55 [FREE Full text] [doi: 10.3389/fphys.2015.00055] [Medline: 25762937]

49. Kim HG, Cheon EJ, Bai DS, Lee YH, Koo BH. Stress and heart rate variability: a meta-analysis and review of the literature. Psychiatry Investig 2018 Mar;15(3):235-245 [FREE Full text] [doi: 10.30773/pi.2017.08.17] [Medline: 29486547]

50. Taye GT, Hwang HJ, Lim KM. Application of a convolutional neural network for predicting the occurrence of ventricular tachyarrhythmia using heart rate variability features. Sci Rep 2020 Apr 21;10(1):6769 [FREE Full text] [doi: 10.1038/s41598-020-63566-8] [Medline: 32317680]

51. Mosley E, Laborde S. A scoping review of heart rate variability in sport and exercise psychology. Int Rev Sport Exerc Psychol 2022 Jul 07:1-75 [FREE Full text] [doi: 10.1080/1750984x.2022.2092884]

52. Shaffer F, Ginsberg JP. An overview of heart rate variability metrics and norms. Front Public Health 2017 Sep 28;5:258 [FREE Full text] [doi: 10.3389/fpubh.2017.00258] [Medline: 29034226]

53. Valensi P, Extramiana F, Lange C, Cailleau M, Haggui A, Maison Blanche P, et al. Influence of blood glucose on heart rate and cardiac autonomic function. The DESIR study. Diabet Med 2011 Apr;28(4):440-449 [FREE Full text] [doi: 10.1111/j.1464-5491.2010.03222.x] [Medline: 21204961]

54. Inamdar A. Correlation between fasting heart rate and fasting plasma glucose level in rural Indians. Eur Heart J 2022 Feb 04;43(Suppl 1):ehab849.158 [FREE Full text] [doi: 10.1093/eurheartj/ehab849.158]

55. Gupta S, Gupta RK, Kulshrestha M, Chaudhary RR. Evaluation of ECG abnormalities in patients with asymptomatic type 2 diabetes mellitus. J Clin Diagn Res 2017 Apr;11(4):OC39-OC41 [FREE Full text] [doi: 10.7860/JCDR/2017/24882.9740] [Medline: 28571189]

56. Nirala N, Periyasamy R, Singh BK, Kumar A. Detection of type-2 diabetes using characteristics of toe photoplethysmogram by applying support vector machine. Biocybern Biomed Eng 2019 Jan;39(1):38-51 [FREE Full text] [doi: 10.1016/j.bbe.2018.09.007]

57. Philip LA, Rajasekaran K, Jothi ES. Continous monitoring of blood glucose using photophlythesmograph signal. In: Proceedings of the 2017 International Conference on Innovations in Electrical, Electronics, Instrumentation and Media Technology. 2017 Presented at: ICEEIMT '17; February 3-4, 2017; Coimbatore, India p. 187-191 URL: https://ieeexplore.ieee.org/document/8116832 [doi: 10.1109/icieeimt.2017.8116832]

58. Moraes JL, Rocha MX, Vasconcelos GG, Vasconcelos Filho JE, de Albuquerque VH, Alexandria AR. Advances in photopletysmography signal analysis for biomedical applications. Sensors (Basel) 2018 Jun 09;18(6):1894 [FREE Full text] [doi: 10.3390/s18061894] [Medline: 29890749]

59. Habbu S, Dale M, Ghongade R. Estimation of blood glucose by non-invasive method using photoplethysmography. Sādhanā 2019 Mar 16;44(6):135 [FREE Full text] [doi: 10.1007/s12046-019-1118-9]

60. Hina A, Nadeem H, Saadeh W. A single LED photoplethysmography-based noninvasive glucose monitoring prototype system. In: Proceedings of the 2019 IEEE International Symposium on Circuits and Systems. 2019 Presented at: ISCAS '19; May 26-29, 2019; Sapporo, Japan p. 1-5 URL: https://ieeexplore.ieee.org/document/8702747 [doi: 10.1109/iscas.2019.8702747]

61. Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol Heart Circ Physiol 2000 Jun;278(6):H2039-H2049 [FREE Full text] [doi: 10.1152/ajpheart.2000.278.6.H2039] [Medline: 10843903]

62. Delgado-Bonal A, Marshak A. Approximate entropy and sample entropy: a comprehensive tutorial. Entropy (Basel) 2019 May 28;21(6):541 [FREE Full text] [doi: 10.3390/e21060541] [Medline: 33267255]

63. Nordström A, Hadrévi J, Olsson T, Franks PW, Nordström P. Higher prevalence of type 2 diabetes in men than in women is associated with differences in visceral fat mass. J Clin Endocrinol Metab 2016 Oct;101(10):3740-3746 [FREE Full text] [doi: 10.1210/jc.2016-1915] [Medline: 27490920]

64. Shi B, Yen SC, Tay A, Tan DM, Chia NS, Au WL. Convolutional neural network for freezing of gait detection leveraging the continuous wavelet transform on lower extremities wearable sensors data. In: Proceedings of the 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society. 2020 Presented at: EMBC '20; July 20-24, 2020; Montreal, QC p. 5410-5415 URL: https://ieeexplore.ieee.org/document/9175687 [doi: 10.1109/embc44109.2020.9175687]

65. Shi B, Tay A, Au WL, Tan DM, Chia NS, Yen SC. Detection of freezing of gait using convolutional neural networks and data from lower limb motion sensors. IEEE Trans Biomed Eng 2022 Jul;69(7):2256-2267 [FREE Full text] [doi: 10.1109/TBME.2022.3140258] [Medline: 34986092]

66. Alwosheel A, van Cranenburgh S, Chorus CG. Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis. J Choice Model 2018 Sep;28:167-182 [FREE Full text] [doi: 10.1016/j.jocm.2018.07.002]

67. Lyssenko V, Jonsson A, Almgren P, Pulizzi N, Isomaa B, Tuomi T, et al. Clinical risk factors, DNA variants, and the development of type 2 diabetes. N Engl J Med 2008 Nov 20;359(21):2220-2232 [FREE Full text] [doi: 10.1056/NEJMoa0801869] [Medline: 19020324]

68. Diabetes risk factors. US Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/diabetes/basics/risk-factors.html [accessed 2023-11-04]

69. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: Proceedings of the 31st Conference on Neural Information Processing Systems. 2017 Presented at: NIPS '17; Dec 4-7, 2017; Long Beach, CA p. 4768-4777 URL: https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf

70. Zhou XH, McClish DK, Obuchowski NA. Statistical Methods in Diagnostic Medicine. Hoboken, NJ: John Wiley & Sons; 2009.

Abbreviations

AI: artificial intelligence
BGEM: blood glucose evaluation and monitoring
CWT: continuous wavelet transform
DM: diabetes mellitus
G-mean: geometric mean
HR: heart rate
HRV: heart rate variability
IFG: impaired fasting glycemia
IGT: impaired glucose tolerance
KTE: Kaiser-Teager energy
ML: machine learning
MSE: multiscale entropy
PNS: parasympathetic nervous system
PPG: photoplethysmography
PSD: power spectral density
SampEn: sample entropy
SHAP: Shapley additive explanations
SNS: sympathetic nervous system
SVM: support vector machine
T1DM: type 1 diabetes mellitus
T2DM: type 2 diabetes mellitus

XSL·FO
RenderX