May – June 2020
ISSN: 0193-4120 Page No. 10165 - 10171

Personalized Non-Invasive Blood Glucose Monitor Using Machine Learning Models

[1]Pradeep Kumar Anand- [1, 4] Electrical and Computer Engineering, Sungkyunkwan University, Suwon 16419, South Korea
[2]Dong Ryeol Shin - [2] College of Software, Sungkyunkwan University, Suwon 16419, South Korea
[3]Mudasar Latif Memon - [3] IBA Community College NaushahroFeroze, Sukkur IBA University, Sindh 65200, Pakistan
[4]Ranjita Kumari- [1, 4] Electrical and Computer Engineering, Sungkyunkwan University, Suwon 16419, South Korea

Article Info
Volume 83
Page Number: 10165 - 10171
Publication Issue:
May - June 2020

Article History
Article Received*: 19 November 2019*
Revised*: 27 January 2020*
Accepted*: 24 February 2020*
Publication*: 18May 2020*

Abstract:
In the past few decades have witnessed immense growth in non-invasive sensing technologies. Unfortunately, all non-invasive blood glucose monitors developed until now lacks to measure glucose value accurately. This paper presents the concept of "personalized non-invasive blood glucose monitor" to measure glucose levels accurately for all diabetic patients using different machine learning models. These models are random forest, support vector machine (SVM), multi-layer perceptron (MLP), decision tree, and adaptive boosting (AdaBoost). Our concept consists of both invasive and non-invasive sensors on a single device. Initially, during clinical trials, several patients' blood glucose is measured both invasively and non-invasively. Then the paired data is divided into five different groups. Five machine learning models are trained for each group having paired data. Each machine learning model predicts non-invasive values accurately based on patients' characteristics. Once, errors in predicted non-invasive values are within the acceptable error range, patient measures blood glucose by non-invasive methods only. Our concept is applied on a baseline simulation data, the MARD is reduced from 36.1% to 12.4% for adaptive boosting machine learning model. The minimum to maximum error is reduced from -221 ~ 55% to -55 ~ 48%.
Keywords*: glucose monitor, blood glucose, diabetes, non-invasive glucose monitor, machine learning.*

INTRODUCTION

As per the world health organization (WHO), approximately 422 million people suffered from diabetes and around 1.6 million people died in the year 2014 [1-2]. The number of diabetic patients is growing every year with a ~8.5% rate and it is expected that by 2030 more than a billion people will be suffering from this disease. WHO has declared diabetes as the number one disease in the world. Diabetes is caused by insulin disorder. The causes of insulin disorder can be by birth or due to an unhealthy diet and less physical activities. These two causes are responsible for Type 1 and Type 2 diabetes [1-2]. Type 1 diabetes (also referred to as Juvenile-onset) is caused due to ineffective production of insulin produced by beta cells of the pancreas [2]. Whereas, Type 2 diabetes (also known as adult-onset) arises from ineffective use of insulin inside the body [2]. The main reason for Type-2 diabetes is a metabolic disorder, due to high blood glucose involving insulin resistance. According to WHO, Type-2 diabetes is responsible for approximately 90% of all diabetes cases. Both Type-1 and Type-2 diabetic patients need diagnostic and regular monitoring to manage their disease [3 – 9]. All these cause the diabetes diagnostic treatment

TEST Engineering & Management

May – June 2020
ISSN: 0193-4120 Page No. 10165 - 10171

market to reach up to $33 billion in 2016. To reduce patient pain, several non-invasive blood glucose measuring devices are developed in the last few decades. Some of the most successful non-invasive blood glucose monitors are developed as Combo Glucometer (COG) by CNOGA [7], GlucoTrack by Integrity application [6] and Wizmi by Wer2b Limited [5]. However, none of these devices meet the 95% accuracy requirement as defined by the FDA and other countries regulation [4].

SOURCE OF INACCURACY IN NON-INVASIVE BLOOD GLUCOSE MEASUREMENT

In, last several decades, most of the research works are around non-invasive sensor technology, improvement in sensor accuracy, algorithm development, and calibration (clinical trials) methods. However, the inaccuracies from other sources of non-invasive devices are not considered very well. Our paper explains the major sources of inaccuracy in non-invasive blood glucose measurement. To analyze the causes of inaccuracy in non-invasive blood glucose measurement in a systematic way, we developed the cause & effect diagram using the Minitab Fishbone concept as shown in Figure 1. As per the fishbone cause & effect diagram, any effect can be categorized into five groups. These groups are personnel, machines, materials, methods, measurements, and the environment. We analyze the causes of blood glucose inaccuracy under every five groups as discussed below.

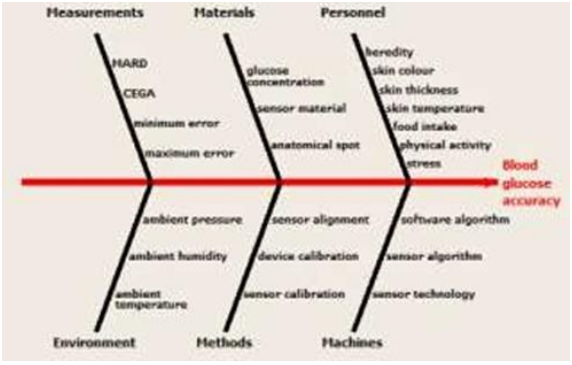

Figure 1. Fishbone cause analysis for error in non-invasive blood glucose measurement

1)Personnel cause: Major sources of glucose inaccuracy are variation within and in between patients. Device accuracyeffects due to rapid change in blood glucose for a patient. Patient food intake, physical activity, stress, and anxiety are the main sources within-patient variation as shown in Figure 1. Blood glucose greatly varies as time elapsed after food intake or physical activity. Patient stress or anxiety also affects blood glucose levels.As glucose concentration in our blood is very low, sensor performance is largely affected due to in between patient's variation. Heredity, skin color, skin thickness [10], and skin temperature are in between patient's variation as shown in Figure 1 under personnel cause. Stem cells, skin color, skin thickness, and body temperature are different for every human. Hence, making a non-invasive blood glucose device, which works accurately for all patients is the biggest challenge. Presently, non-invasive blood glucose monitors (COG, Wizmi, GlucoTrack) do not address inaccuracy due to the causes mentioned. These companies perform clinical trials to address this issue. However, clinical trials on a set of patients partially address these causes of inaccuracy. Our research explained in this paper address these causes for blood glucose inaccuracy by personalizing the non-invasive device for each patient.

2)Machines: In the machine group, we identify sensortechnology, sensor algorithm, and software algorithm as major causes for blood glucose inaccuracy as shown in Figure

1. Non-invasive blood glucose accuracy widely varies based on sensor technology. Sensor and software algorithms play a very important role during detection, the conversation of sensor output into electrical parameters, transforming electrical analog signals to digital values, software data processing, calculation, and presenting data to display. All these are sources of inaccuracy in the

TEST Engineering & Management

May – June 2020
ISSN: 0193-4120 Page No. 10165 - 10171

measurement. These sources of inaccuracy are addressed during design and development.

Material: In the material group of Figure 1, we identify patient blood glucose concentration greatly affects the non-invasive measurement accuracy as glucose concentration inside the human body is very low. Degradation in sensor material over the use also affects the ability to read blood glucose. Anatomical spot (finger, ear lobe) plays a very important role for blood glucose measurement. Blood glucose concentration range inside the human body is fixed and we cannot do anything about this. This is a constraint and we have to develop a non-invasive blood glucose monitor under this constraint. Sensor degradation is addressed by performing reliability and life test. Anatomical spot selection is addressed during product design by careful selection of body parts with more blood circulation, less skin thickness, and a patient's comfort.

1)Method: Some of the causes for inaccuracy in blood

glucose measurement are sensor calibration, device calibration, and sensor alignment [11] to a non-invasive device detector as shown in Figure 1 under method cause. Sensor and device calibrations are accomplished during clinical trials. The patient measurement spot (anatomical region) and its alignment with the sensor play a very crucial role in device accuracy [11]. This is addressed by patient training.

2)Measurement: Non-invasive device accuracy is measured by mean average relative difference (MARD) or by

Clarke error grid analysis (CEGA), and error range. Error range represents the minimum and maximum percentage error for a patient.

3)Environment: Ambient temperature, humidity, and pressure may also affect sensor accuracy based on its technology. During development, functional and performance test is performed to ensure that the device is working well under room ambient condition.

Based on fishbone cause & effect analysis, it is clear that existing non-invasive blood glucose monitors (COG, Wizmi, GlucoTrack) have taken care of most of these causes listed under the machine, material, method, measurement, and environment during device development, clinical trials, and patient training. However, none of these non-invasive blood glucose monitors have addressed the causes of error in the personal category completely. Our research described in this paper proposes the non-invasive blood glucose monitor concept by personalizing the device for a patient using machine learning to address these personal causes as well.

PERSONALIZED GLUCOSE MONITOR

We propose the concept of personalized blood glucose monitor as shown in Figure 2. It has both invasive and non-invasive sensors on a single device. As both sensors are available on a single device, it gives the advantage to check the initial error (E1) in non-invasive blood glucose value (X) with respect to invasive or reference value (R) in real-time during the clinical trial and personalized calibration as shown in Figure 2. Machine learning software builds various models based on invasive (R) and non-invasive value (X) with the goal to predict non-invasive value (Y) equal to invasive or reference value (R). The final error (E2) is calculated based on non-invasive predicted (Y) and reference value (R) as shown in Figure 2. All these invasive, non-invasive, non-invasive predicted, initial errors and final error are shown on a display. A mode select button helps the user to choose invasive or non-invasive blood glucose measurement method. Table 1 defines each parameter referred to in this paper.

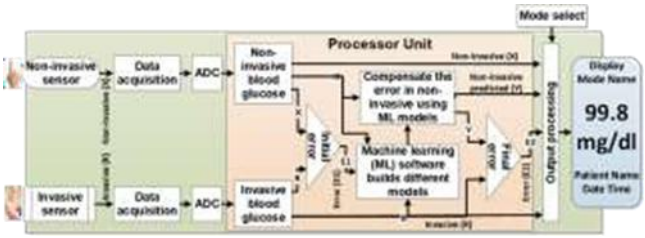

Figure 2. Personalized non-invasive blood glucose monitor

May – June 2020
ISSN: 0193-4120 Page No. 10165 - 10171

Table 1. Parameter Definition.

Parameter	Definition
R	Invasive (reference) glucose value
X	Non-invasive glucose value
Y	Non-invasive predicted glucose value
E1	Initial error in non-invasive glucose value with respect to a reference
E2	Final error in non-invasive predicted glucose value with respect to a reference

Initially, personalized non-invasive blood glucose monitor is trained during clinical trials. For each patient duringclinical trials, we measure blood glucose both invasively (R) and non-invasively (X). Initial error (E1) is calculated for each pair of data (X and R) as shown in Figure 2. Several paired data are collected during clinical trials. A machine learning software is written in python. The purpose of the machine learning software is to train the following listed models based on the paired data collected during clinical trials. The machine learning models are random forest, support vector, multilayer perceptron, decision tree, and adaptive boosting (AdaBoost). Training to test ratio is set at 70:30 for each model.

- Random forest
- Support vector machine (SVM)
- Multilayer Perceptron (MLP)
- Decision tree
- Adaptive boosting (AdaBoost)

The output of machine learning models is to predict non-invasive blood glucose value (Y) accurately with respect to invasive or reference value (R) as demonstrated in Figure 2. To do this, the final error (E2) is calculated based on the non-invasive predicted value (Y) and reference (R). The goal is to train each model based on the least mean average relative difference (MARD) and root mean square error (RMSE) considering in between patient variations (heredity, skin color, skin thickness, skin temperature). Various hyperparameters are changed to optimize each machine learning models. Once training is completed the test is performed on 30% of data collected during the clinical trial. Both training and test data are randomly selected during each run to reduce the biasing error. During this step, we identify the best performing model, save the optimized hyperparameters and rank them in the order. Once machine learning models are successfully built during clinical trials, then personalized glucose monitor device is re-calibrated based on individual patient behaviors (heredity, skin color, skin thickness, skin temperature, food habit, physical activity, stress level). In this step, the device is used by a patient. Both invasive (R) and non-invasive (X) values are measured for a patient and paired data are collected. Each machine learning model trained during step 1 clinical trials is retrained based on data collected for a patient in this step. Retraining of the models is done based on priority order identified during step 1 as the number of data are limited in this step for a patient. Once the device starts predicting non-invasive value (Y) accurately based on updated machine learning models within the pre-defined accuracy limit then personalized calibration stage completes as shown in Figure1. At this point, blood glucose is measured non-invasively bythe best performing model and the device becomes personalized for a user. The best performing model with their hyperparameters is saved for that particular patient in the personalized calibration library.

SIMULATION DATA AND GROUPS FORMATION

To validate the concept of a personalized glucosemonitor, we choose GlucoTrack as a reference device. Integrity application GlucoTrack uses three different types of sensors based on ultrasonic, electromagnetic and thermal technology tomonitornon-

May – June 2020
ISSN: 0193-4120 Page No. 10165 - 10171

invasivebloodglucoselevels[6].GlucoTrackhas performed a clinical trial on 91 subjects and achieved meanabsoluterelativedifferences(MARD)as23.4%[6] .To prove our concept, we use GlucoTrack invasive and non-invasive data as a baseline. We reproduced simulation data within +/-1% of accuracy as shown in Table 2. Reproduced simulation baseline data has mean and median ARDas23.9%and16.4%comparedto23.4%[6]and16.5 % [6], respectively for GlucoTrack data. Minimum and maximum ARD or error is -221% and 61% for both GlucoTrack and reproduced baseline data as shown in Table 2.

Table 2. Reproduced GlucoTrack data.

Parameters	Unit	Glucotrack [6]	Reproduced baseline
Mean ARD	%	23.4	23.9
Median ARD	%	16.5	16.4
Minimum Error	%	-221	-221
Maximum Error	%	61	61
Invasive Range	mg/dl	65 ~ 492	65 ~ 492
Non-invasive range	mg/dl	80 ~ 352	80 ~ 352

Next, we develop the groups based on patient condition and blood glucose values. For random blood glucose checks, patients with a glucose level of 50 – 80 mg/dl are considered in hypoglycemia condition. Patients with 81 – 115 mg/dl glucose level are considered as non-diabetes and 116 – 150 mg/dl are considered as a pre-diabetes state. Patients with glucose levels range from 151 - 180 mg/dl are considered as hyperglycemia and >181 mg/dl is considered a high diabetes condition. Our approach divided the entire reproduced baseline paired data set into 5 groups.

TEST RESULT

We trained all five machine learning models, random forest, support vector machine, multilayer perceptron, decision tree, and AdaBoost separately in every five groups using 70% baseline data. We kept the same test data for each machine learning models to have an unbiased comparison. After optimizing the hyperparameters, we performed the test on 30% data set selected randomly and present the test results in Table 3. Table 3 represents overall performance after combining the result from five groups. It is very evident that all models performed well and successfully predicted the non-invasive glucose values accurately and hence reduce the final error (E2) significantly. For the random forest, the final MARD reduces from 23.0 to 11.6% on 30% of test data. Similarly, final MARD reduces from 23.1 to 13.5% for the SVM, 19% to 24.9% for the MLP, 24.9% to 13.0% to the decision tree, and 36.1% to 12.4% for the AdaBoost. MLP has the least reduction in final MARD, while AdaBoost has the highest reduction. Usually, MLP performs well on large data. It is very clear that AdaBoost is the best performing model as minimum error reduces drastically from -221% to -55% and maximum error reduces from 60% to 48%. We choose the AdaBoost model for personalized glucose monitor as MARD is reduced from 36.1% to 12.4% and minimum to maximum errors are reduced from -221% ~ 60% to -55% ~ 48%.

Table 3. Result of personalized glucose monitor machine learning models based on five groups.

Model	Minimum		Maximum		MARD	
	Initial	Final	Initial	Final	Initial	Final
Random Forest	-152	-64	60	52	23.0	11.6
SVM	-149	-60	54	60	23.1	13.5
MLP	-136	-66	52	58	19	24.9
Decision Tree	-221	-64	60	58	24.9	13.0
AdaBoost	-221	-55	60	48	36.1	12.4

Figure 3 shows the graphical representation of the test result of glucose value for ―invasive (R)‖, ―non-invasive (X)‖ and ―non-invasive predicted by AdaBoost (Y)‖. The non-invasive curve clearly shows that it has a wide variation with respect to invasive or reference blood glucose. However, we can clearly interpret that the non-

May – June 2020
ISSN: 0193-4120 Page No. 10165 - 10171

invasive predicted curve is following invasive (reference) glucose curve very well.

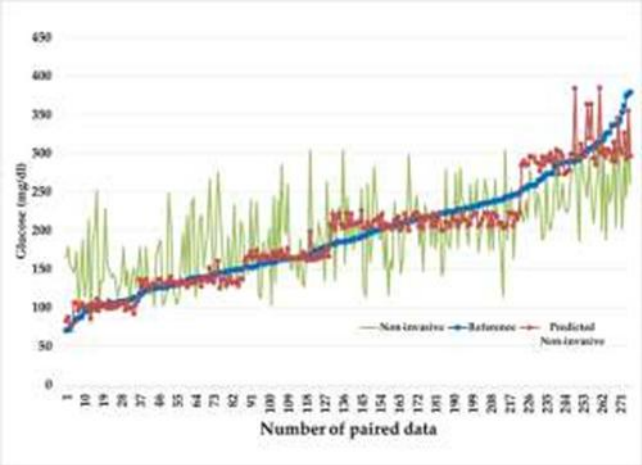

Figure 3. Non-invasive measured, non-invasive predicted with respect to the reference.

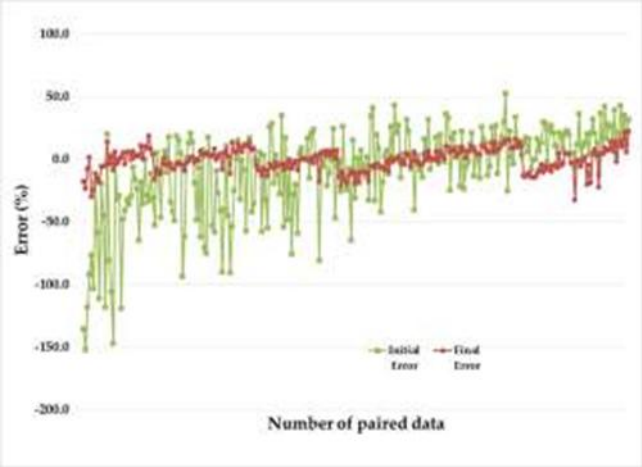

Figure 4. Percentage error in non-invasive easured and predicted blood glucose value.

Figure 4 shows the graphical representation of initial error (E1) in baseline compared to final error (E2) in personalized glucose monitor using AdaBoost. From the plot in Figure 4, it is very evident that baseline data has a very wide error in non-invasive blood glucose ranging from -221 ~ 55% compared to an error in personalized monitor using AdaBoost ranging from -55 ~ 48%.

CONCLUSION

We proposed a personalized glucose monitor using five different machine learning models. Our approach efficiently predicted non-invasive values based on patients' characteristics and reduced the error in non-invasive blood glucose significantly. We divided the paired invasive and non-invasive blood glucose data into five groups based on patient diabetes conditions named hypoglycemia, normal, pre-diabetes,hyperglycemia,andhighlydiabetes.Wetrained five different machine learning models named as random forest, support vector, multilayer perceptron, decision tree, andAdaBoostforeachofthefivegroups.Resultsshowedthat the AdaBoost predicts non-invasive values more accurately than other machine learning models. The personalized glucose monitor with AdaBoost applied on a baseline simulation data consist of invasive and non-invasive paired data. It reduced the MARD from 36.1% on baseline data to 12.4%forthepersonalizedmonitorwithAdaBoost.Theerror range (minimum to maximum) reduced from -221 ~ 55% for baseline data to -55 ~ 48% for the personalizedmonitor.

References

1. Global Report on Diabetes, World Health Organization. Available online:
2. https://www.who.int/diabetes/global-report/en/ (accessed on 25 November 2019).
3. Makaram, P.; Owens, D.; Aceros, J. Trends in nanomaterial-based non-invasive diabetes sensing technologies. Diagnostics. 2014, 4(2), pp.27-46.
4. Yilmaz, T.; Foster, R.; Hao, Y. Radio-frequency and microwave techniques for non-invasive measurement of blood glucose levels. Diagnostics. 2019, 9(1), p.6.
5. Cappon, G.; Acciaroli, G.; Vettoretti, M.; Facchinetti, A.; Sparacino, G. Wearable continuous glucose monitoring sensors: a revolution in diabetes treatment. Electronics. 2017, 6(3), p.65.
6. Villena Gonzales, W.; Mobashsher, A.T.; Abbosh, A. The progress of glucose monitoring—A review of invasive to minimally and non-invasive techniques, devices and sensors. Sensors. 2019, 19(4), p.800.

May – June 2020
ISSN: 0193-4120 Page No. 10165 - 10171

7. Harman-Boehm, I.; Gal, A.; Raykhman, A.M.; Naidis, E.; Mayzel, Y. Noninvasive glucose monitoring: increasing accuracy by combination of multi-technology and multi-sensors. Journal of diabetes science and technology. 2010, 4(3), pp.583-595.
8. Segman, Y. Device and method for noninvasive glucose assessment. Journal of diabetes science and technology. 2018, 12(6), pp.1159-1168.
9. Malchoff, C.D.; Shoukri, K.; Landau, J.I.; Buchert, J.M. A novel noninvasive blood glucose monitor. diabetes care. 2002, 25(12), pp.2268-2275.
10. Ali, H.; Bensaali, F.; Jaber, F. Novel approach to non-invasive blood glucose monitoring based on transmittance and refraction of visible laser light. IEEE Access. 2017, 5, pp.9163-9174.
11. Turgul, V.; Kale, I. Simulating the effects of skin thickness and fingerprints to highlight problems with non-invasive RF blood glucose sensing from fingertips. IEEE Sensors Journal. 2017, 17(22), pp.7553-7560.
12. Turgul, V.; Kale, I. Influence of fingerprints and finger positioning on accuracy of RF blood glucose measurement from fingertips. Electronics Letters. 2017, 53(4), pp.218-220.