Available online at www.garmian.edu.krd

Passer Journal

Passer 6 (Issue 1) (2024) 58-73

http://passer.garmian.edu.krd/

Development of a Non-Invasive Blood Glucose Monitoring Device Using Machine Learning Technology

Makarand Mohan Jadhav[1*], Avinash N Sarwade[2], Vijay M Sardar[3], Hemlata M Jadhav[4]

[1]*Dept. of Information Technology, NBN Sinhgad Technical Institute Campus, Pune, Maharashtra, India*
[2]*Electrical Engineering Department Sinhgad College of Engineering, Pune, Maharashtra, India*
[3] *Jayawant Shikshan Prasarak Mandal (JSPM), Pune, Maharashtra, India*
[4] *Marathwada Mitra Mandal's College of Engineering (MMCOE), Pune, Maharashtra, India*

Received 04 March 2023; revised 19 April 2023;
accepted 01 November 2023; available online 17 January 2024

DOI: 10.24271/PSR.2023.388351.1272

ABSTRACT

Background: Blood glucose monitors are critical to diabetes management. There is no permanent medicine to cure diabetes. Presently, invasive glucose blood meters extract blood sample by pecking a needle into the patient's fingers. This results in the formation of copious calluses on the fingertips and causes more pain to lure blood again and again for repetitive measurements.
Objectives: The study aims to develop Non-Invasive Glucometer to monitor the glucose level of a person using a Wi-Fi module. A variation in amplitudes, and phases of received packets, helps to measure glucose levels. A Hampel filter is used to suppress abrupt amplitude variations occurring due to environmental effects. Further, the Fast-Tree Regression algorithm is used to train the model for different glucose concentrations for accurate prediction and detection of diabetes. It also reduces dataset dimension for minimizing the training time of the device. Thereafter, Clarke Error Grid Analysis helps to estimate the accuracy.
Materials and Methods: Two ESP32 Wi-Fi devices, are installed on a computer for real time sensing of Channel State Information (CSI) between an Receiver Access Point and Transmitter Station. Further, additional header information such as MAC address, RSSI, and other metadata along with the CSI is sent for all 64 subcarriers. Here, statistical regression analysis is considered only to confirm the results.
Results: The accuracy achieved is 95 % with coefficient of determination in terms of an R^2 value of 0.99. The device measures glucose level in less than 3 sec. It can store 2000 test results with time, date, daily and weekly average reports for random, before, and after the meal. A containers containing air and 5% Glucose solution helps to validate the models behavior with specific glucose content.
Conclusion: The portable painless device is found to be useful to monitor the glucose level at home and office. The benefits is low-cost and Non –Invasive.

Keywords: AP, CSI, CSS, EGA, MAE, MSE, ML, RMSE, RSSI, STA.

1. Introduction

Blood glucose monitors are critical to diabetes management. Presently, there is no permanent medicine to cure diabetes. Patients with diabetic symptoms are kept within the range by controlling levels of blood glucose. Various techniques are available to measure and monitor the regulation of glucose. Presently, blood meters use glucose oxidase and platinum-coated strips for glucose measurement, where, initially blood is pecked from the patient's finger and placed on enzyme strips. Further, amperometry analysis is carried out on yield hydrogen peroxide engendered from the chemical reactions of glucose and oxygen.

* Corresponding author
E-mail address: makj123@yahoo.com (Instructor).
Peer-reviewed under the responsibility of the University of Garmian.

In the invasive glucose monitoring method, a blood sample was extracted by pecking a needle into the patient's fingers found to be a painful process. This results in the formation of copious calluses on the fingertips and causes more pain to lure blood again and again for repetitive measurements. Whereas, Non-invasive glucose monitoring techniques have been heavily researched over the past several decades[1].

Diabetes mellitus is affecting the Indian population seriously in the present day. It is classified as type 1 diabetes. The other is type 2 diabetes. A glucose concentration in the range of 72-134 represents a non-diabetic patient. Whereas, the range 111-330 mg/dL represents a diabetic patient. This diabetes damages the eyes, kidneys, and nerves. Further, it protectorates hypertension due to kidney failure[2]. Patients with diabetic symptoms are kept within the range by controlling levels of blood glucose. Various techniques are available to measure and monitor the regulation of glucose. Hence, designing blood glucose measuring is perilous to controlling and managing diabetes. In this paper, the design, and development of a reliable low-cost non-invasive glucose monitoring system are presented. It is developed using the 802.11a Wi-Fi module. This module measures glucose levels based on the permittivity changes in human blood. The module has 64 sub-carrier frequencies spaced with 312.5 kHz bandwidth. The total generated 64 amplitudes and phases are used for channel estimation and in turn to measure blood glucose levels. A Hampel filter suppresses abrupt amplitude variations occurring due to the propagation effects. To reduce the dimension of the dataset some key frequency data is used to train the model for different glucose concentrations using machine learning for accurate detection of diabetes. It also reduces dataset dimension and minimizes the time for training the system. The performance or error metric is crucial to ensuring the success of solving a real-world problem with machine learning. Below we discuss the following types of performance metrics for fast forest regression. The first performance metric used is a mean absolute error. The actual values in the data set are considered. Thereafter, predicted values are also stored in the data set. The average of the absolute difference values helps to measure the average of the residuals in the set of data. The second performance metric considered is a mean squared error. The average of the squared difference of data value is measured to find the variance of the residuals. The last metric considered is the root mean squared error. It is calculated as the square root of the mean squared error. It finds residuals' standard deviation. The coefficient of determination is called R-squared. A linear regression model characterizes the variance proportion in the dependent variable. The R-square is a scale-free score. Its value is always less than one. The lesser Mean Absolute Error (MAE), Mean Square Error (MSE), as well as Root Mean Square Error (RMSE) values, infers better regression model accuracy[3]. On the other hand, a high R-square value is desirable.

This section elaborates on work done in the implementation of a non-invasive device. The methodology and their experimentation are divided into five parts based on Diabetes Mellitus, Glucose Monitoring, Non-Invasive Glucose Monitoring techniques, Machine Learning, and Wi-Fi Channel State Information. In[5], glucose contents in blood with high blood pressure cause obesity. Further, patients with more cholesterol range are also suffering from cardio as well as microvascular diseases. The pervasiveness of obesity in diabetes patients is 54.8%. Indirect expenses ensuing diabetes in[6] is $39.8. In[7] advised events like mealtimes, exercise, as well as hypoglycemia symptoms throughout the day, are recorded. The data gained helps to recommend a diabetes schedule to maintain glucose levels. Investment to design and developing portable non-invasive devices is capturing the market as presented in[8]. In[9] a method to extract glucose by applying a potential to the skin is explained. Further, an electrochemical enzymatic device measures glucose levels. This technique creates mild skin irritation. Glucose absorptions in human tissue affect micro-circulation and are explored in[10] without considering tissue-blood dynamics. A deep artificial neural network using ultra-wide frequencies is presented to estimate glucose levels in the blood[11]. It is a non-invasive method. The accuracy achieved is 88%. Here, no sample of blood is used to monitor the diabetics. A real-time smartwatch is proposed to monitor glucose levels. A chipless tag passive sensor is used in[12]. An interstitial fluid helps to measure variation in glucose concentrations. A principle of optical detection is proposed to measure the glucose level in the blood. It is known as the near-infrared spectroscopic method[13]. The precision of the device is analyzed with models in the library of neural networks. A principle of reflection and absorption is implemented in[14] to measure glucose content in the blood. An optical source with 940 nm as well as 1300 nm is considered. The device is corroborated and verified by three human beings. First is a healthy patient. Second, tested with the pre-diabetic patient. At the last, also tested with diabetic patients. A regression model is proposed to check the precision of the device. A bio antenna with an artificial neural network is presented as a non-invasive system. A user-friendly interface is also provided to monitor and measure the glucose level and health grade of the patient[11]. The accuracy achieved is 90.6%. A variation in material permittivity to measure physiological deviations is expressed in[15] with microwave frequency. This resonant frequency is captured with the resonator device. Machine knowledge framework is arrayed to train the system and predict the glucose level in the blood[16, 17]. A probability distribution function to understand the features, as well as labels, is considered in[18] to predict glucose levels. It is termed a forest algorithm. Prediction accuracy in the range of 10 -30 % is presented. An optimization technique to predict the speed with a machine learning model is proposed. Robustness, as well as competence, is verified in various data collections[19].

Wi-Fi device information in[20] provides signal strength at the device end. It recognizes various human activities. This analysis is used to measure the glucose level of diabetic patients. A radial basis function is used to detect variation in the amplitude of a used Wi-Fi device[21]. It is known as the Wi-Fi system. It has three components. The first is the sensor module. A pre-processor is the second component. The last is the neural network detection module. Liquids have diverse dielectric constant values. It results in signal amplitude variation. Therefore, sub-carriers in Wi-Fi devices considered in[22] has different frequency characteristics. It is observed that a single packet provides all the channel information. It is used to measure glucose levels and maintain the health of diabetic patients. A complete analysis of Wi-Fi technology is presented in[23]. A channel state information obtained from this device is used to detect glucose contents in the blood. It has explored three applications. The first is for signal detection. The second is used to recognize objects. The third is used to estimate the parameters of the signal. A principle of change in permittivity and conductivity of the material is considered in[24]. It provides variation in amplitude as well as phase of the Wi-Fi signals. This variation is used to measure glucose levels. Its efficacy in the Internet of Things is presented as a case study. A system to detect ripeness in fruit is proposed in[25]. A Wi-Fi device available at home is used. Such physiological variation is used to measure the glucose level of diabetic patients. A channel State Information obtained from Wi-Fi is considered for the covid framework[26]. It is carried out in four steps. First signal amplitude is measured. The unwanted signal outliers are removed using a Hampel filter. The principal component analysis is carried out to extract frequency components. In the end, covid symptoms are detected. The same technique is extended to measure the glucose level in the blood to monitor and detect diabetics. Thus, a portable painless, and cost-effective non-invasive glucose monitoring device is developed to monitor the glucose level of a person at home as well as at the office using a Wi-Fi module.

In 2020, it is estimated that the number of people diagnosed with diabetes could rise to over 17 million, costing an estimated $192 billion around the globe. This will result in chronic conditions such as vascular disease, renal complications, and a variety of neurological symptoms that lead to more mortality. Hence, an attempt to reduce the cost of treating diabetes and the pain involved in extracting blood from finger strips.

The main objective of this work is to study the effectiveness of glucose monitoring using Channel State Information (CSI) signal amplitudes. The range of blood glucose levels that can be measured is in the range of 40 mg/dl to 400 mg/dl. The proposed device takes less than 3 sec to measure blood glucose levels. The device can store 2000 Blood Glucose Test Results with Time, Date, and daily as well as weekly Average reports for Random or before or After Meals. The device provides USB interface connectivity and a large display with a 3.5-inch LCD to the patients. The input of the device is the patient hand between the antennas. The output of the device is blood Glucose Value in mg/dL. The operating conditions for the device are 40 C to 450 C temperature with a humidity level of less than 98%. The altitude range in which the device can be used is 0 to 3094 meters. The prediction engine used in this paper is for predicting results and it returns the results as measured glucose levels in the blood (mg/dL).

Here the CSI Data is considered to train & measure the Blood Glucose Levels by keeping a Human Hand between STA (Transmitting Station) and AP (Receiving Access Point). The variations in Blood Glucose Levels affect the amplitudes of the subcarriers. The system is trained with multiple individuals by taking actual Blood Glucose Levels & training the system using ML.Net Fast Forest Quantile Regression Method. For checking the Glucose Level from the system the ML.Net Model Predicts the Glucose Level. Machine Learning (ML) can be explained as finding results by comparing historical data that is stored using an algorithm (Model). The process begins with collecting data, labeling or marking data, training data to generate a Model, and consuming the Model to predict the result. Evaluation of the Model is done by comparing the measured versus predicted results.

Imagine diabetics who no longer have to suffer the frequent pain of drawing blood to measure their blood sugar levels. Where monitoring is performed more frequently to gain better disease management, less suffering, fewer medical complications, and lowered health care costs. In day-to-day life, an embarrassing approach of pricking blood from the human body is used to measure and record the glucose levels of diabetic patients. This process is carried out either in the clinic or at home. However, obtained glucose levels in the blood are at that specific time only. Hence there is a need to monitor and measure unobserved glycemic changes in the blood several times a day to avoid complications. Therefore, designing painless and low-cost non-invasive meters for continuous glucose measurement is always been areas of interest. Such personal device is found better for glucose control and taking necessary preventive actions at the right time to save life[26].

However, available glucose meter offered to the customer has their disadvantages. One of these meter types needs a blood sample extracted by pecking a needle into the patient's fingers. This is a painful process. This results in the formation of copious calluses on the fingertips and causes more pain to lure blood again and again for repetitive measurements. The second type, the continuous glucose meter provides incessant monitoring of glucose levels based on measuring interstitial fluid. This meter requires added calibration to blood samples and is not cost-effective. The third type of glucose meter is comparatively cheap and uses disposal electrode strips. However, strips are expensive for the patient who requires to monitor and measure glucose levels several times a day. Presently there are two types of glucometers available in the market. One is based on an invasive method and the other on a non-invasive method. Both methods measure the glucose level in the blood cells.

Imagine diabetics who no longer have to suffer the frequent pain of drawing blood to measure their blood sugar levels. Where monitoring is performed more frequently to gain better disease management, less suffering, fewer medical complications, and lowered health care costs. In day-to-day life, an embarrassing approach of pricking blood from the human body is used to measure and record the glucose levels of diabetic patients. This process is carried out either in the clinic or at home. However, obtained glucose levels in the blood are at that specific time only. Hence there is a need to monitor and measure unobserved glycemic changes in the blood several times a day to avoid complications. Therefore, designing painless and low-cost non-invasive meters for continuous glucose measurement is always been areas of interest. Such personal device is found better for glucose control and taking necessary preventive actions at the right time to save life. Further, ML-based studies are mostly restricted to data with high signal-to-noise ratios. In this work, a long-established Random Forest algorithm is modified to take into account uncertainties in measurements as features and in assigned classes as labels.

This paper consists of six sections. Section 2 investigated the work on the implementation painless non-invasive blood glucose measuring device. Section 3 presents the Machine Learning technique used to train the Glucose Prediction Model. Whereas, Section 4 depicts the results and findings of the Non-Invasive Glucometer. Section 5 presents the conclusion and discusses the results followed by references. The future work that can be extended by the researchers is also described.

2. Materials and method

2.1 Implementation painless non-invasive blood glucose measuring device

Imagine diabetics who no longer have to suffer the frequent pain of drawing blood to measure their blood sugar levels. Where monitoring is performed more frequently to gain better disease management, less suffering, fewer medical complications, and lowered health care costs. In day-to-day life, an embarrassing approach of pricking blood from the human body is used to measure and record the glucose levels of diabetic patients. This process is carried out either in the clinic or at home. However, obtained glucose levels in the blood are at that specific time only.

Hence there is a need to monitor and measure unobserved glycemic changes in the blood several times a day to avoid complications. Therefore, designing painless and low-cost non-invasive meters for continuous glucose measurement is always been an area of interest. Such a personal device is found better for glucose control and taking necessary preventive actions at the right time to save a life. Figure 1 represents the block diagram of the proposed non-invasive glucometer. The channel estimation signal response transmitted from the Wi-Fi Transmitter to the Receiver while transferring a data packet is analyzed for the dielectric properties of the sample (hand) kept between the transmitter and receiver. A dielectric characterization is used to estimate the dielectric property of the sample with low-cost commodity Wi-Fi devices. The channel state information is captured and the amplitude values for respective sub-carriers are computed[27]. To reduce the dimension of the dataset some key frequency data is taken. A system setup is shown in Figure 2 with detailed specifications.

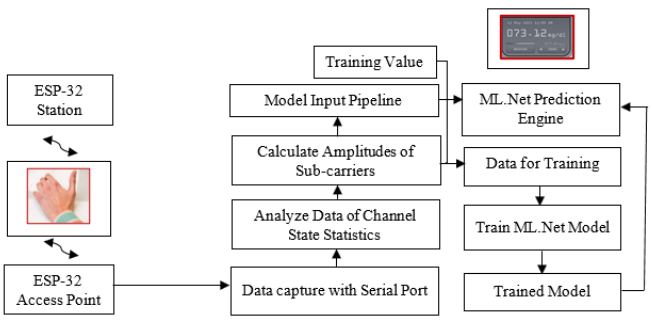

Figure 1: Block Diagram of Proposed Non-Invasive Glucometer.

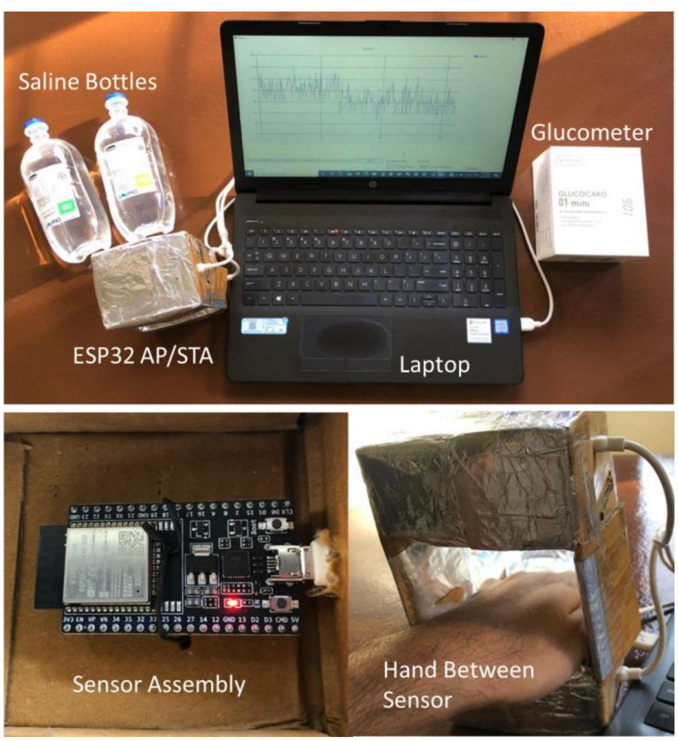

Figure 2: Experimental Set-up of Proposed Non-Invasive Glucometer.

Specifications of Painless Non-Invasive Glucometer Device

Application: To determine blood glucose in whole blood by a Non-Invasive method

Operating Conditions: Temperature: 4^0 C to 45^0 C;

Humidity: < 98%

Operating Altitude: 0 to 3094 meters.

Display: Large Display with 3.5-inch LCD

Measuring range: 40 mg/dl to 400 mg/dl

Measuring time: less than 3 seconds

Memory Capacity: 2000 Blood Glucose Test Results with Time, Date, and Daily as well as weekly Avg report for Random/ before/ after Meals.

Connectivity: USB Interface

The model is trained using ML.Net Fast Forest Regression. Using the reduced dataset, the Model is trained for different glucose concentrations. Using a water-soluble glucose solution, the accuracy obtained is around 90%. Further, real patient data is to be collected & the machine-learning approach is used to train the system based on the data values collected. The description of a Non-Invasive Glucometer to measure Glucose Levels in Human Blood using WiFi CSI Data & Machine Learning is described in Algorithm. In practice signals propagating between source and destination are affected due to the physical properties of the wireless medium. Further, resource allocation at the physical layer is done based on RSSI as well as CSI values obtained from Wi-Fi devices. Received Signal Strength Indication (RSSI) provides information on wireless channel properties and packet delivery status. On the other hand, CSI gets amplitude and phase information of subcarriers used for communication. All this information is used to estimate the channel properties of the communication link. We have used the CSI Data to train & measure the Blood Glucose Levels keeping a Human Hand between STA and AP. The variations in dielectric properties of Blood Glucose affect the amplitudes for the subcarriers. The system is trained with multiple individuals by taking actual Blood Glucose Levels & training the system using ML.Net Fast Forest Quantile Regression Method. For checking the Glucose Level from the system, the ML.Net Model predicts the Glucose Level. A lightweight ESP-32 low-cost Channel State Statistics (CSS) data assortment tool is proposed and deployed. It is a standalone unit to maintain a Wi-Fi sensing system. It is preferred exclusively for large-scale schemes. Espressif IoT Development Framework in C is instigated to develop a code base for ESP-32. The code base consists of two specific applications. It runs on ESP-32 controller pairs. The first application is an active AP. It sets an onboard Wi-Fi stack. This allows devices to connect by making and receiving requests. Whereas, the second application is an active station. It routinely connects to sending requests to the server. The server is running on the access point. The devices automatically initiate communication with this active application. The user-level application collects CSS data it sends to a serial port of the computer for processing. The proposed CSS tool can collect and record data without the help of external devices. Thus, arrayed in enormous quantities even in space-limited situations. Further, can be attached to mobile stuff due to its lightweight design. This tool accesses all 64 subcarriers. The resolution achieved is 8 bits. It is on par with available tools.

Here ESP32 based CSI data collection tool is proposed which is lightweight, can work standalone, and can be easily deployed at a low cost. This provides an opportunity to build more practical and easy-to-maintain Wi-Fi sensing systems, especially for large-scale systems. Further, the codebase for ESP32s is implemented using the Espressif IoT Development Framework (ESP-IDF) in C. The codebase consists of two specific applications which can be run on a pair of ESP32 microcontrollers. The first application is an active access point which initializes the onboard WiFi stack to allow devices to connect, make requests, and receive requests. The second application is an active station which automatically connects to the AP, then sends requests to the server running on the AP. With both applications active, it is possible for the devices to automatically initiate communication but more importantly, it is possible for our user-level application to collect CSI data for further processing. CSI data is automatically written to a serial port of the Computer.

With the proposed ESP32 CSI tool, the ESP32 can collect and record CSI directly without requiring the functionality of any external devices. This in conjunction with the small size of 5cm × 3cm and weight less than 10g of the ESP32 compared to a desktop computer or even a laptop or smartphone means that the proposed tool is very agile and can easily be deployed even in space limited environments in massive amounts, can be repositioned easily and can be attached to mobile objects without much burden thanks to its lightweight design. The tool also gives access to 64 subcarriers where the resolution of each imaginary and real number is 8 bits which is on par with other tools.

Two ESP32 Wi-Fi devices are installed on a personal computer. These are utilized to emit and receive signals. To perform real-time sensing in a given environment, we consider offloading tasks between an ESP32-RX as AP and ESP32-TX as STA . Specifically, as shown in Fig. 3, the ESP32-RX receives raw CSI from TX.

When collecting CSI from the ESP32-AP to the computer, serial communication throughput must also be considered when viewing the RX rate. Further, additional header information per CSI frame such as MAC address, RSSI, and other metadata along with the channel state information for all 64 subcarriers is send resulting in each frame of size 8kb.

Figure 4 represents the test setup configuration to measure the glucose level in the blood. Two ESP32 Wi-Fi devices are installed on a personal computer. These devices emit as well as receive information signals. To accomplish real-time sensing in a specified scenario, offloading tasks amongst an ESP-32RX as AP and ESP-32TX as STA are considered in this paper. The ESP32-RX receives raw CSI from TX. When collecting CSI from the ESP32-AP to the computer, serial communication throughput is considered as a QoS parameter to view the receive rate. Here, extra data per CSS frame is added as a header for all 64 subcarriers. It consists of physical address, RSSI, as well as channel state statistics. The generated frame is of size 1kbytes. The baud rate selection demarcates connection speed for communication. It is observed that the 115.2 kbits/s baud rate achieves maximum throughput. It can retain serial consistency. The logged CSS data divulges features of the subcarrier's channel response. The employed device measures CSS data. It creates a set of 32 transmissions. Further, captures 64 subcarriers' CSS data. Here, a bottle consisting of various materials is kept in the propagation path. It impacts the received signal. In the end, the channel state also varies consequently. ML.Net is used to create a model to depict measurable signal characteristics as well as material dielectric properties.

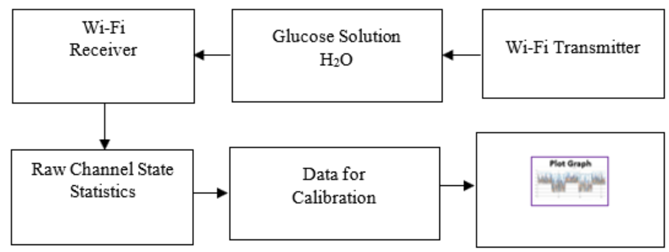

Figure 3: Test Setup of Blood Glucose measuring device.

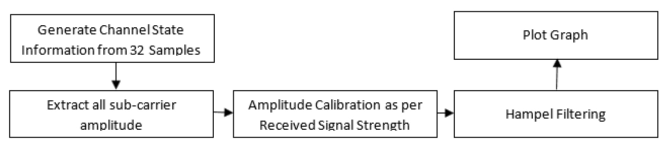

Figure 4: Process of CSI Data extraction to Plot Amplitude Graph.

2.2 CSI Estimation Model

The multipath components' amplitude, as well as phase, are arriving at the receiver. The signal communicates expressed with equation (1)

$$A_r e^{j\emptyset_r} = A_{LOS} e^{j\emptyset_{LOS}} + A_m e^{j\emptyset_m} \tag{1}$$

Here, $A_m e^{j\emptyset_m}$ epitomizes the multi-path effect of signals. The amplitude gets attenuated relative to the dielectric constant and conductivity of the dispersive medium. Equation (2) represents received signals $R(f)$in the frequency domain.

$$R(f) = S(f) * H(f) \tag{2}$$

Here, $S(f)$ represents a transmitted signal. Whereas, the $H(f)$channel response measured by the access point is expressed with $H(f)$AP. It is discovered as the Channel State Statistics data.

2.3 Channel State Statistics

CSS is a metric used in the OFDM system. It describes amplitude as well as phase among subcarrier frequencies. This carries transmitted signals to the receiver. A set of pilot symbols detect these variations across subcarriers. These pilot symbols are interleaved in the transmitted message. This preamble estimates vector H describing CSS given by equation (3)

$$y^{(i)} = H^{(i)}x^{(i)} + n^{(i)} \quad (3)$$

Where $y^{(i)}$specifies the detected signal at the receiver. Further,$x^{(i)}$ expresses pilot symbols. Whereas, $n^{(i)}$ represents the noise vector. CSS of each subcarrier is a complex number. It consists of both imaginary and real parts of vector H. Thereafter, the real and imaginary parts of the individual subcarrier are combined. It provides subcarrier amplitude $A^{(i)}$ and phase $\emptyset^{(i)}$ as given by equation (4)

$$A^{(i)} = \sqrt{\left(h_{im}^{(i)}\right)^2 + \left(h_r^{(i)}\right)^2} \text{and } \emptyset^{(i)} = atan2(h_{im}^{(i)}, h_r^{(i)}) \quad (4)$$

In this work, the transmitter and receiver are kept line-of-sight. However, it is observed that measurements obtained anterior of line of sight by two meters reduced prediction accuracy from 92.1% to 83.1%. The positioning of ESP32 devices is a perilous factor to improve prediction accuracy. In this paper, newly developed lightweight as well as stand-alone tool is explored. Thus, the amplitude depends upon the received signal strength that is captured by the parameter RSSI. The collected CSI frame is calibrated as per the RSSI value to minimize the effect of the automatic gain control of the transmitted signal. The real-world datasets have missing data and anomalies. A Hampel filter is used to identify outliers and replaces them with representative values by calculating the median and estimated window's standard deviation σ. For any point in the window, if it exceeds 3σ out from the window's median, then the Hampel filter identifies the point as an outlier and replaces it with the window's median. Figure 3 explains the process of CSI Data extraction to plot the amplitude graph.

Algorithm: To determine blood glucose in whole blood by a Non-Invasive method

----

Operating Conditions: Temperature: 4C to 45 C and Humidity: <98%

Operating Altitude: 0 to 3094 meters.

Display: Large Display with 3.5-inch LCD

Measuring range: 40 mg/dl to 400 mg/dl

Measuring time: less than 3 seconds

Memory Capacity: 2000 Blood Glucose Test Results with Time, Date, and daily as well as weekly average report for Random or Before or After Meals.

Connectivity: USB Interface

Input: 32 samples of CSI Data for sub-carrier no 8,16,24,32,40,48,56 out of 64

Output: Blood Glucose Level in mg/dL

Steps :

- The ESP32 WiFi STA generates a request in a loop to ESP32 WiFi AP.
- The AP prints the CSI Data to Serial Port
- The Desktop Application reads the CSI Data over Serial Port
- Stores the CSI Data for 32 requests made by STA to AP in 2D_CSI_Array
- The Amplitude is computed for each subcarrier from the CSI Data
- The Amplitude time-series data for 32 instances is stored for the sub-carriers

  8,16,24,32,40,48,56

- Hampel filter is used to remove the outliners in the Amplitude time-series

Algorithm a: Hampel Filter to remove the Outliners in CSI Amplitudes

Definitions

list_amplitudes: time-series data

k: window size

nSigma: threshold

index, y

Program

initialize index as 0

for each item x in list_amplitudes

initialize y as 0

if the index is less or equal to window size k: First Samples

y = k

if index + k greater than length of list_amplitudes : Last Samples

y = -k

index = index + y

stddev = standard deviation of array values from (index – k) to (index + k)

median = median of array values from (index – k) to (index + k)

index = index – y

if the Absolute value of (list_amplitudes[index] – median) > (nSigma x stddev)

list_amplitudes[index] = median

index incremented by 1

loop

return list_amplitudes

• Blood Glucose reading is obtained by regular invasive method

• The obtained Amplitude time series is tagged by the Blood Glucose reading

• A set of 50 Amplitude time series are captured and stored as Training Data

• Training is done for Blood Glucose in 3 step process. One before lunch. Second, an hour after lunch. Third, two hours after lunch.

• The Training is done using ML.Net Fast Forest Regression

Algorithm b: Train Model with ML.Net

Definitions

TRAIN_DATA_FILEPATH

MODEL_FILEPATH

mlContext object of ML.Net

Program

Load training data from TRAIN_DATA_FILEPATH to ML.Net & generate training data-View

Build training pipeline from ml. Context

Train Model mlModel with ML.Net:Regression Training [FastForest]

Evaluate the quality of the model & Print Metrics for the Regression Model

Save Model to MODEL_FILEPATH

• A random reading is taken by capturing Amplitude time series A-Test

• ML.Net uses A-Test to predict the glucose level in the blood

• The Predicted Blood Glucose Level is displayed on the Screen

Algorithm c: Predict Result with ML.Net for given input CSI_DATA

Definitions

PredictionEngine

MODEL_FILEPATH

mlModel

Input

Result

Program

CreatePredictionEngine :

Load mlModel from MODEL_FILEPATH

Use PredictionEngine generated from ML.Net CreatePredictionEngine

PredictionEngine.Predict :

use the Predict function to return the Result for Input

• The system can be re-trained by inputting correct Blood Glucose readings by invasive methods.

----

3. Machine Learning Techniques

This technique finds result by comparing historical data that is stored using an algorithm or model. The process begins with collecting data, labeling or marking data, training data to generate a model, and consuming the model to predict results. Here, evaluation of the model is done by comparing the measured versus predicted results.

With ML.NET, you can create and train your own ML models, which makes ML.NET very flexible and adaptable to your specific data and business domain scenarios. ML.NET follows the same basic steps for nearly every scenario; it combines data loading, transformations, and model training to create machine learning models. ML Context is the starting point for all ML.NET operations. The ML Context is used for all aspects of creating and consuming an ML.NET model. Once you have an instance of an ML Context, you can load and transform data, choose the best algorithm for your machine learning task, and train your model. Once trained, model can be tested for accuracy, save it to disk, and use it to make predictions. Further, it uses training data to find patterns to make predictions on new and unknown data. The inputs for learning are called features, which are the attributes used to make predictions. Whereas, output of learning is called the Label, which is the actual prediction.

The Fast Forest Regression module in ML.Net Machine Learning is an implementation of random forest quantile regression using decision trees. Random forests can be helpful to avoid overfitting that can occur with decision trees. A decision tree is a binary tree-like flow chart, where at every interior node, one decides which of the two child nodes to continue to, based on the value of one of the features of the input. Fast Forest Quantile regression is useful if you want to understand more about the distribution of the predicted value, rather than get a single mean prediction value. This method has many applications, including Predicting prices, Estimating student performance or applying growth charts to assess child development, and discovering predictive relationships in cases where there is only a weak relationship between variables. This regression algorithm is a supervised learning method, which means it requires a tagged dataset that includes a label column. Because it is a regression algorithm, the label column must contain only numerical values.

3.1 ML.net

In ML.NET each scenario that desires to be evaluated follows pre-defined steps. It starts with combining data loading, applying a transform, and performing training to create machine learning models. Further, ML. context is the initial point for all this operation. It creates and consumes all aspects of this .NET segment. Thus, an instance of ML Context created helps to load as well as transform data. Thereafter, helps to select the suitable task to train the proposed scenario. Once trained, the model is tested for accuracy and the sets are saved to disk. It is used to make estimates of diabetics. Machine learning uses these trained data to discover patterns. This pattern helps to improve prediction accuracy on new as well as unknown data. The inputs are called Features. This attribute is used for predictions. Whereas, the outputs are called Labels. This attribute is the real predicted value.ML.NET has thirty algorithms for different tasks. Thus, choosing a precise task for the scenario is the biggest challenge for the researchers. In this paper, Fast Forest Regression trainers that aptly fit the proposed application are chosen. The data is applied to an estimator. It learns from this data to create a transformation. For the model training process, input is training data. Whereas, the output is a trained model. In general, this trained model transform feature into prediction. It is stored as an integrated binary file. In the end, loaded into .NET applications. In line with this, a prediction engine is considered a suitable API to make single predictions.

3.2 Fast Forest Regression Trainer

The regression analysis process estimates the relationships between the input variables (criteria) and one or more independent results (predictions). It is a supervised learning method. That needs a ticketed dataset. Thus, the label column in this algorithm contains only numerical values. It explains the changes in input variables to changes in select results. The conditional expectation of the input variables based on results in terms of the average value of the trained variables is given at the time the input variables are changed. The Fast Forest Regression segment in ML.Net machine learning implements a random forest quantile regression algorithm using decision trees. It avoids over-fitting that occurs using decision trees. It is a binary tree flow process. At all interior nodes, a decision to select a child node is done based on the input features. Thus, it is observed that Fast Forest Quantile regression provides a distribution of the predicted value. The following types of performance metrics are used for regression analysis given by Equations 5 - 8. Let N be the number of samples, y_i be measured and $\hat{y}$ predicted value. A set consisting of actual as well as predicted data is prepared. Measures the absolute variance in the data set given by equation (5)

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}| \tag{5}$$

MSE calculates the residual change given by equation (6)

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2 \tag{6}$$

RMSE that represents the standard aberration of residuals is given by equation (7)

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y})^2} \tag{7}$$

The R-squared value that represents the variance quantity in the dependent variable is given by equation (8)

$$R^2 = 1 - \frac{\sum(y_i - \hat{y})^2}{\sum(y_i - \bar{y})^2} \tag{8}$$

4. Results and Discussion

The glucose level in the blood is measured using a non-invasive glucometer base record, with Fasting, One and Half hours after eating and helping in diagnosing the diabetics. For experimentation, the CSI Data is considered to train & measure the Blood Glucose Levels keeping a human hand between STA and AP. The variations in glucose levels affect the amplitudes of the subcarriers. The system is trained with multiple individuals by taking actual glucose Levels & training the system using ML.Net Fast Forest Quantile Regression Method. For checking the glucose level from the system the ML.Net Model predicts the glucose Level for Diabetics detection.

Different methods to diagnose diabetes are:

- Fasting plasma glucose test: This test is best done in the morning after an eight-hour fast (nothing to eat or drink except sips of water).
- Random plasma glucose test: This test can be done at any time without the need to fast.
- A1C test: This test, also called HBA1C or glycated hemoglobin test, provides your average blood glucose level over the past two to three months. This test measures the amount of glucose attached to hemoglobin, the protein in your red blood cells that carries oxygen.
- Oral glucose tolerance test: In this test, the blood glucose level is first measured after an overnight fast. Then you drink a sugary drink. Your blood glucose level is then checked at hours one, two, and three.
- Gestational diabetes tests: There are two blood glucose tests if you are pregnant. With a glucose challenge test, you drink a sugary liquid and your glucose level is checked one hour later. You don't need to fast before this test. If this test shows a higher-than-normal level of glucose (over 140 ml/dL), an oral glucose tolerance test will follow (as described above).

Further, the Data Set considered for experimentation are: 1 mg/dL represents Air , 91 mg/dL represents the glucose value in the morning without eating anything, 162 mg/dL represents One and a Half hours after Breakfast glucose value, 89 mg/dL represents the glucose value of a second person without eating anything and 110 mg/dL represents one and a Half Hours after breakfast glucose value of the Second Person.

The Clarke Error Grid Analysis (EGA) was developed in 1987 to quantify the clinical accuracy of patient estimates of their current blood glucose as compared to the blood glucose value obtained in their meter. It was then used to quantify the clinical accuracy of blood glucose estimates generated by meters as compared to a reference value. Eventually, the EGA became accepted as one of the "gold standards" for determining the accuracy of blood glucose meters. The performance or error metric is crucial to ensuring the success of solving a real-world problem with machine learning. Below we discuss the following types of performance metrics for fast forest regression :

MAE: The Mean absolute error represents the average of the absolute difference between the actual and predicted values in the dataset. It measures the average of the residuals in the dataset.

MSE: Mean Squared Error represents the average of the squared difference between the original and predicted values in the data set. It measures the variance of the residuals.

RMSE: Root Mean Squared Error is the square root of Mean Squared error. It measures the standard deviation of residuals

R-Squared: The coefficient of determination or R-squared represents the proportion of the variance in the dependent variable which is explained by the linear regression model. It is a scale-free score i.e. irrespective of the values being small or large, the value of R square will be less than one.

Further, to validate the model's behavior during the presence of a liquid with specific glucose content, an experiment wherein the containers containg air and 5% Glucose solution are placed between AP at receiver and STA at transmitter. The antenna is kept at the LOS Line of Sight.

The above graph shows the graph amplitudes for all Subcarriers. When the container is empty the amplitude received is depicted for all the Subcarriers. Further, it is observed that the amplitude is significantly attenuated when a 5% Glucose solution is placed between AP & STA. It is seen from the experimental results that the amplitude of the real subcarrier has high variations compared to the modeled amplitude as represented in different colors. This is caused by variation of the diminished reflection of the reflected path due to movements in the environment. Here, such attenuation is assumed constant over time.

The other discrepancy between the model and the actual result is the phase shift which is a result of a combination of factors such as the initial phase of the unoccupied room, reflections, and the fluctuation in human movement velocity. To minimize the phase shift, the distance between AP & STA is reduced. It is observed that amplitude attenuation (k) occurs with different concentrations of Glucose liquid placed between AP & STA.

Figure 5 represents the flow chart of the proposed Low Cost and Reliable Machine Learning based Painless Non-Invasive Blood Glucose Measuring Device.

(a) (b)

(a) Slide in Left palm and Measure Glucose level (b) Glucose level is displayed in 2-3 Sec (c) Train Glucometer for accurate results.

4.1 Captured Amplitude Time-series

Figure 5 shows the graph of amplitudes for all Subcarriers. It is assumed that an amplitude attenuation (k) will occur with different concentrations of glucose liquid placed between AP and STA. It is observed that when the container is empty the amplitude received is depicted for all the Subcarriers. The amplitude is significantly attenuated when a 5% Glucose solution is placed between AP and STA. The subcarrier has high variations in amplitude. It is more than the modeled amplitude. The reason for this variation is the occurrence of the reflected path. This results in suppressed amplitude due to environmental movements. This attenuation is kept persistent over the period. The additional incongruity is the phase shift. It is measured amid the prototype and obtained results. The occurrence is due to the consolidation of three factors. First is the original phase of the unoccupied room. The second is reflections and the third is variation in movement velocity of patients. To abate the phase shift, the distance between AP and STA is reduced.

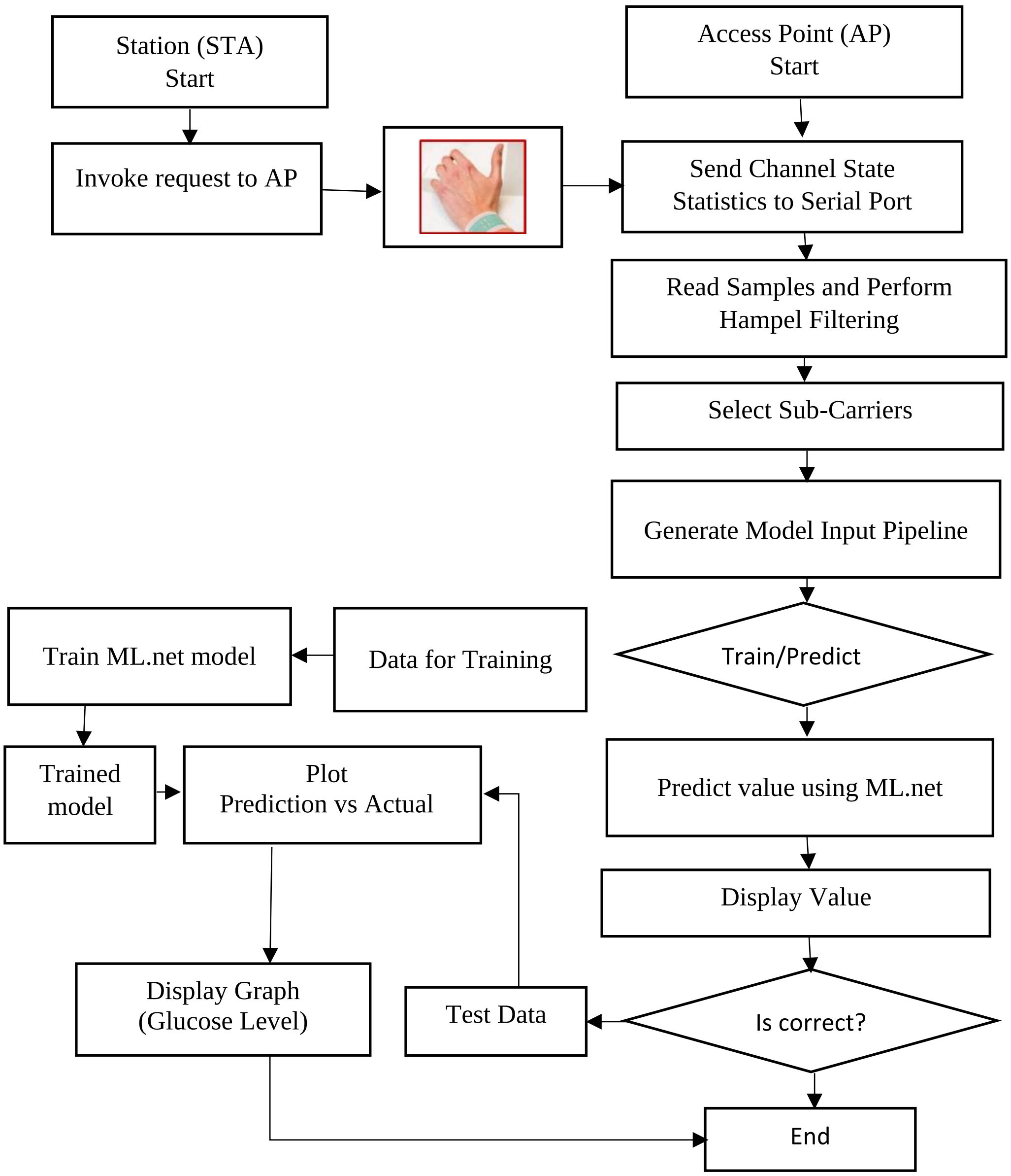

Figure 5: Flow Chart for Non-Invasive Glucometer.

It is assumed that an amplitude attenuation (k) will occur with different concentrations of glucose liquid placed between AP and STA. Figure 6 shows the graph of amplitudes for all Subcarriers. In Figure 7, an experiment for three containers containing Honey, DNS, and 5% Dextrose solution is carried out. The amplitude for Honey is highly attenuated than for 5% Dextrose.

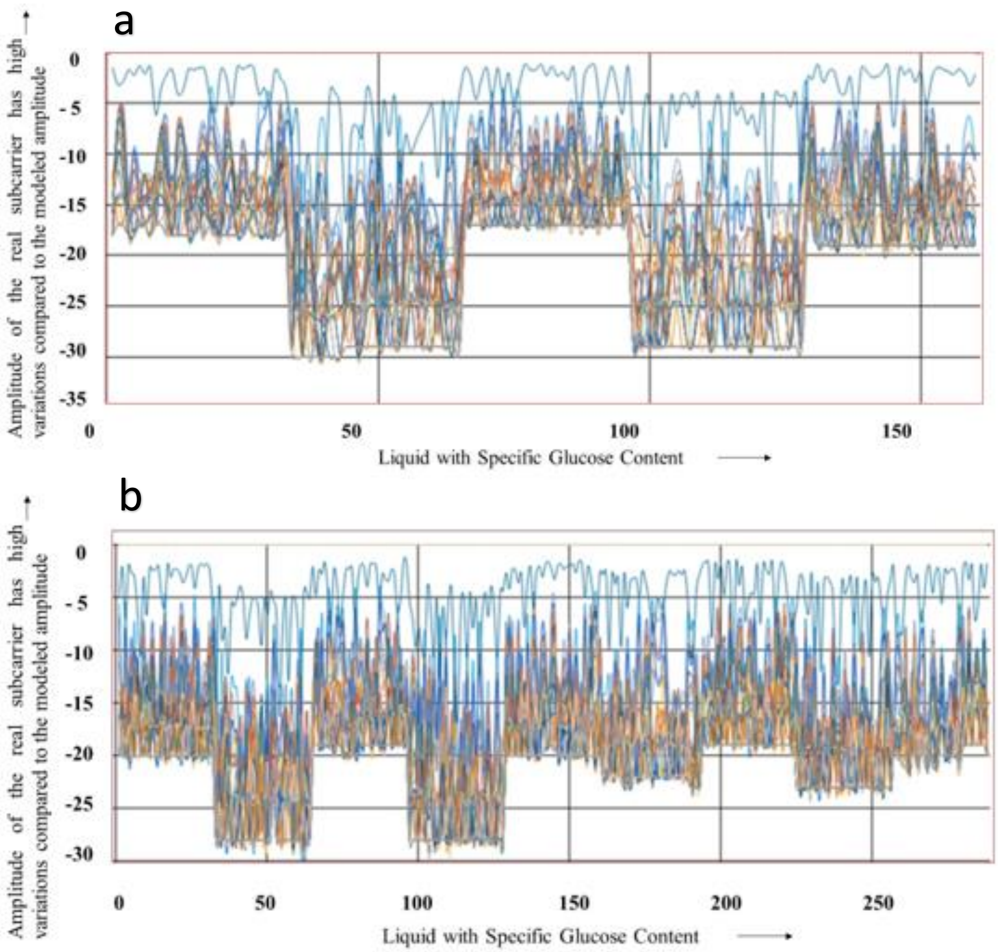

Figure 6: a) CSI Amplitude Graph of Different Glucose Solutions, b) CSI Amplitude Graph of Air & 5% Dextrose Bottle.

DNS and 5% Dextrose show similar attenuation since both solutions contain the same amount of glucose. The next experiment that is performed is to measure blood glucose using an available invasive glucometer at a different time of day. The experiment was carried out at three time instants such as before lunch, one hour after lunch when the glucose level is maximum, and two hours after lunch when the glucose level falls back to normal as shown in Figures 8 and 9.

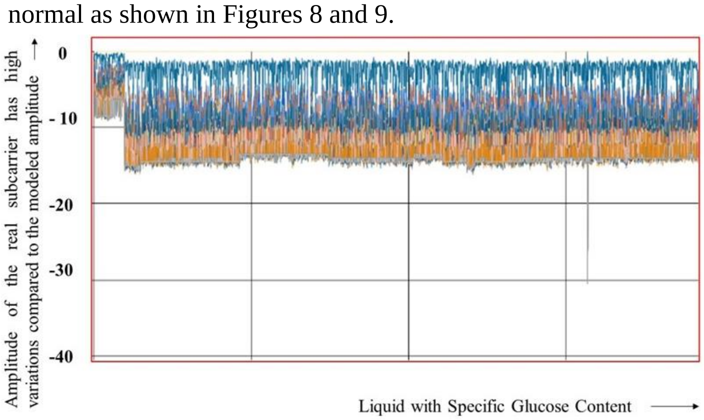

Figure 7: Graph for Blood Glucose Measured 162 mg/dL (One hours after Lunch).

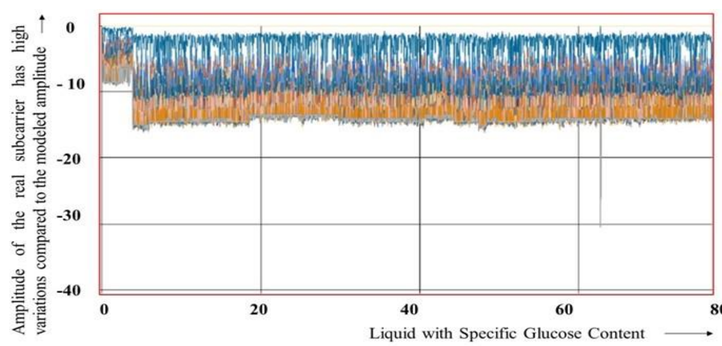

Figure 8: Graph for Blood Glucose Measured 162 mg/dL (Two hours after Lunch).

4.2 Device Accuracy Estimation

EGA technique is presented to quantify the clinical precision of blood glucose value estimation. It helps to determine blood glucose device accuracy.An attempt to minimise the estimation error obtained from existing and device glucose values is as shown in Figure 10. The Simulation parameters considered are given in Table 1. Whereas Table 2 explores the significance of Fast Forest Regression to measure performance metrics.

CSI is a metric used in OFDM for describing amplitude and phase variations across multiple subcarrier frequencies as wireless signals are transmitted between a transmitter and receiver. To detect these variations across subcarriers, OFDM systems transmit a set of known shared pilot symbols which are then used to estimate a vector H that is describe as CSI

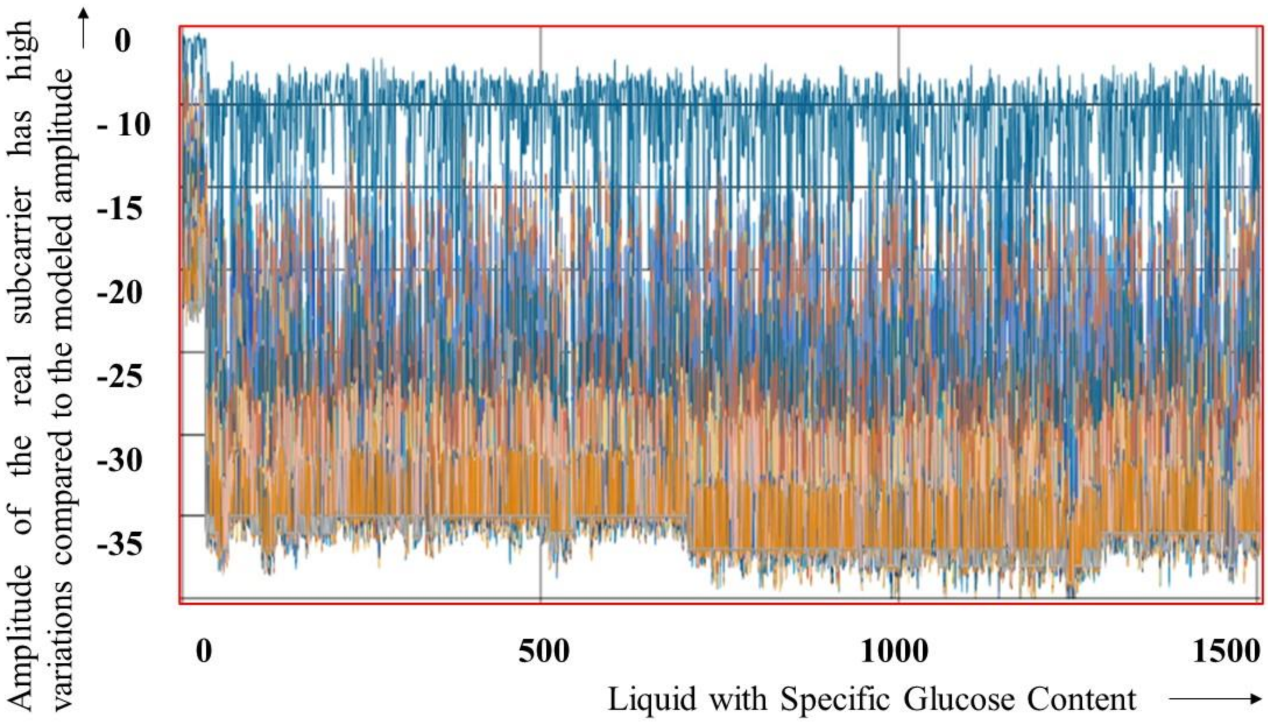

Figure 9: Graph for Blood Glucose Measured 162 mg/dL (Two hours after Lunch).

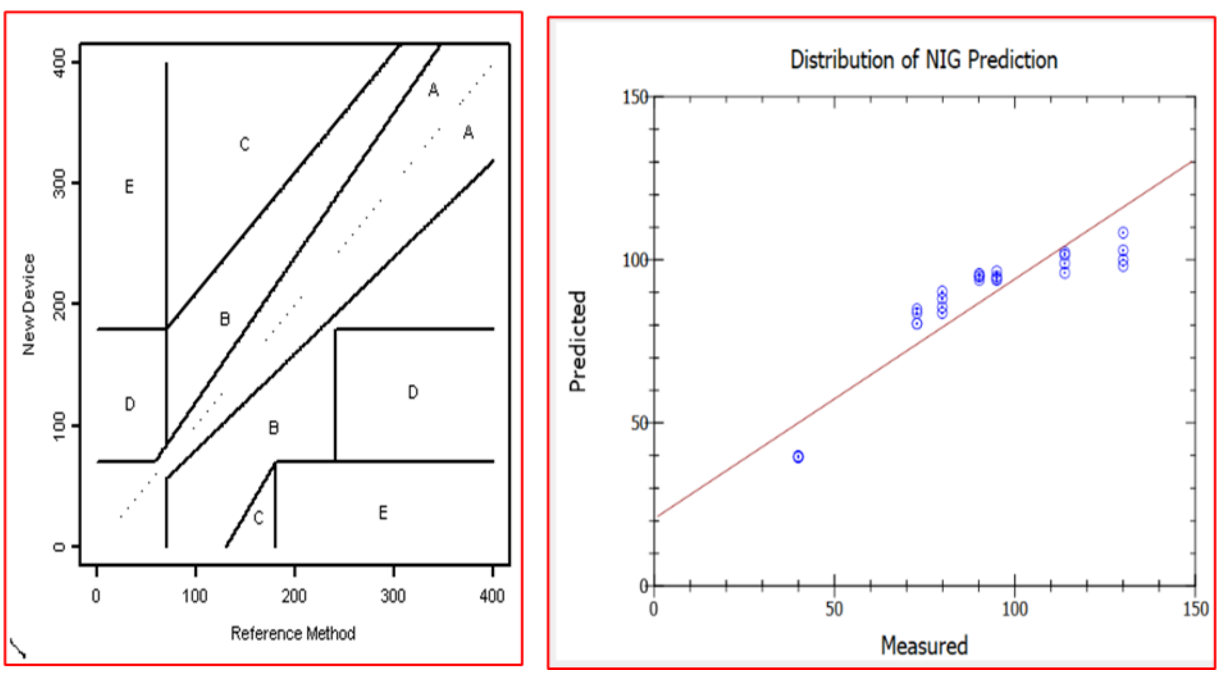

Figure 10: Clarke Error Grid & Prediction Distribution.

Table 1: Simulation Parameters.

CSI Data of 64 sub-carriers: 20 MHz

No of Samples collected: 32

Operating Frequency: Min 2412 MHz and Max 2484 MHz

Sub-carriers chosen for data extraction are from Channels: 8,16,24,32,40,48,56

Corresponding Frequency: 2424,2432,2440,2448, 2456,2464 and 2464 MHz

Table 2: Fast Forest Regression to measure performance metrics.

Sample No	Measured mg/dL	Predicted mg/dL	Difference	Square of Difference	Square of Mean Difference
1	1	1	0	0	11071.2484
2	1	1	0	0	11071.2484
3	1	1	0	0	11071.2484
4	91	92.00062	1.00062	1.001240384	202.1907676
5	91	95.655716	4.655716	21.67569147	111.6040964
6	91	97.07973	6.07973	36.96311687	83.54453567
7	91	95.34177	4.34177	18.85096673	118.3358879
8	91	89.399826	1.600174	2.56055683	282.9182534
9	91	93.89663	2.89663	8.390465357	151.8654482
10	91	96.778465	5.778465	33.39065776	89.14258316
11	91	90.618164	0.381836	0.145798731	243.4172866
12	91	87.04726	3.95274	15.62415351	367.5939591
13	91	91.592804	0.592804	0.351416582	213.9548628
14	91	90.77502	0.22498	0.050616	238.5474072
15	91	94.6395	3.6395	13.24596025	134.1079803
16	91	97.26509	6.26509	39.25135271	80.19041311
17	91	92.25208	1.25208	1.567704326	195.1027891
18	91	91.28411	0.28411	0.080718492	223.0808101
19	162	168.99992	6.99992	48.99888001	3941.318355
20	162	163.47804	1.47804	2.184602242	3278.483145
21	162	169.6749	7.6749	58.90409001	4026.524334
22	162	160.90439	1.09561	1.200361272	2990.38251
23	162	155.16806	6.83194	46.67540416	2395.912578
24	162	166.26031	4.26031	18.1502413	3604.838825
25	162	157.75468	4.24532	18.0227419	2655.823243
26	162	164.308395	2.308395	5.328687476	3374.261634
27	162	164.17213	2.17213	4.718148737	3358.449372
28	162	163.6561	1.6561	2.74266721	3298.905583
29	162	166.9136	4.9136	24.14346496	3683.713081
30	162	163.38966	1.38966	1.931154916	3268.370025
31	162	160.24478	1.75522	3.080797248	2918.676854
32	162	162.04095	0.04095	0.001676903	3115.978459
33	162	164.2525	2.2525	5.07375625	3367.771056
34	89	83.068825	5.931175	35.17883688	535.9769039
35	89	92.29897	3.29897	10.88320306	193.7950763
36	89	94.70739	5.70739	32.57430061	132.540189
37	89	95.65782	6.65782	44.32656715	111.5596464
38	89	86.677334	2.322666	5.394777348	381.9157944

39	89	90.446274	1.446274	2.091708483	248.8104319
40	89	93.68226	4.68226	21.92355871	157.1949243
41	89	87.857353	1.142647	1.305642167	337.1868048
42	89	91.34395	2.34395	5.494101603	221.2968636
43	89	89.42226	0.42226	0.178303508	282.1640691
44	89	90.94326	1.94326	3.776259428	233.378785
45	89	83.76677	5.23323	27.38669623	504.1475374
46	89	89.21464	0.21464	0.04607033	289.1822687
47	89	90.22177	1.22177	1.492721933	255.9433631
48	89	91.137146	2.137146	4.567393025	227.4924848
49	89	91.274925	2.274925	5.175283756	223.3552668
50	89	90.55	1.55	2.4025	245.5489
	Mean Value	**MAE**	**MSE**	**RMSE**	**R-Squared**
	106.22	2.81102446	12.7701003	3.57352771	0.99289051

Simulation Parameters

CSI Data of 64 sub-carriers: 20 MHz 802.11n/ac channel consists of 64 subcarriers

No of Samples collected: 32

Operating Frequency: Min 2412 MHz, Max 2484 MHz

The sub-carriers was chosen for data extraction:

Channels: 8,16,24,32,40,48,56

Corresponding Frequency: 2424,2432,2440,2448,2456,2464,2464 MHz

----

Algorithm: Performance Evaluation of proposed Glucose Monitoring Device

----

Input: 32 samples of CSI Data for sub-carrier no 8,16,24,32,40,48,56 out of 64

Output: Blood Glucose Level in mg/dL

Steps:

The ESP32 WiFi station (STA) generates a request in a loop to ESP32 WiFi Access Point (AP)

The AP prints the CSI Data to Serial Port

The Desktop Application reads the CSI Data over Serial Port

Stores the CSI Data for 32 requests made by STA to AP in 2D_CSI_Array

The Amplitude is computed for each subcarrier from the CSI Data

The Amplitude time-series data for 32 instances is stored for the sub-carriers 8,16,24,32,40,48,56

Hampel filter is used to remove the outliners in the Amplitude time-series

Blood Glucose reading is obtained by regular invasive method

The obtained Amplitude time-series is tagged by the Blood Glucose reading

A set of 50 Amplitude time-series are captured and stored as Training Data

Training is done for Blood Glucose before lunch, one hour after lunch & two hours after lunch.

The Training is done using ML.Net FastForest Regression

A random reading is taken by capturing Amplitude time-series A-Test

ML.Net uses A-Test to predict the Blood Glucose Level

The Predicted Blood Glucose Level is Displayed on the Screen

The system can be re-trained by inputting correct Blood Glucose readings by invasive methods.

Various error metric is used to determine the R^2 value. However, the R^2 value obtained is 0.99. that results in accuracy of 95 %. Further detailed analysis like ANOVA, Full factorial Analysis is the future scope of this study. Therefore, the work is limited only to the effectiveness of Non-Invasive Blood Glucose Measuring devices using Wi-Fi modules. Further, statistical analysis is considered only to confirm the results and not to consider the mathematical modeling of the current system in this paper.

A comparison Table is added in the new manuscript on page no.

Conclusion

A glucose concentration in the range of 111-330 mg/dL represents a diabetic patient. A blood glucose meter is used most at home to measure glucose content in the blood as per the doctor's prescription. A realistic reusable low-cost non-invasive glucose monitoring system is developed using an 802.11a Wi-Fi module. A variation in amplitudes, as well as phases of received packets, helps to measure blood glucose levels. A Hampel filter is used to suppress abrupt amplitude variations occurring due to environmental effects. A total of three containers are placed between the transmitter and receiver containing air, water, and 5% glucose solution to record variations in sub-carrier amplitudes. Further, the Fast-Tree Regression algorithm is used as one machine learning technique to train the model for different glucose concentrations for accurate prediction and detection of diabetes. It also helps to reduce dataset dimensions for minimizing the time for training the system. The R^2 value obtained is 0.99 with achieved accuracy of 95 %. The portable painless and cost-effective non-invasive glucose monitoring device is found to be useful to monitor the glucose level of a person at home as well as at the office using a Wi-Fi module. The benefits of the device is low-cost and Non –Invasive.

Conflict of interests

The authors confirm that there is No conflict of interest to declare for this publication.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors would like to thank the editor and anonymous reviewers for their comments that help improve the quality of this work.

References

1. Islam, M., Ali, M. S., Shoumy N. J., Khatun S., Karim M. S. A., & Bari B. S., Non-invasive blood glucose concentration level estimation accuracy using ultra-wide band and artificial intelligence. *SN Applied Sciences*, **2(2)**, 1-9, doi:10.1007/s42452-019-1884-3 (2020).
2. Baghelani M., Abbasi Z., Daneshmand M., & Light P. E., Non-invasive continuous-time glucose monitoring system using a chipless printable sensor based on split ring microwave resonators. *Scientific Reports*, **10(1)**, 1-15 doi:10.1038/s41598-020-69547-1 (2020).
3. Jain P., Joshi A. M., Agrawal N. & Mohanty S, iGLU 2.0: A new non-invasive, accurate serum glucometer for smart healthcare. *arXiv preprint arXiv*:2001.09182, https://doi.org/10.48550/arXiv.2001.09182 (2020).
4. Islam, M., Min, C. L., Shoumy, N. J., Ali, M. S., Khatun, S., Karim, M. S. A., & Bari, B. S., Software module development for non-invasive blood glucose measurement using an ultra-wide band and machine learning. *In Journal of Physics, IOP Publishing Conference Series*, **1529(5)**, 052-066, doi:10.1088/1742-6596/1529/5/052066 (2020).
5. Ahmed, Z., Amizadeh S., Bilenko M., Carr, R., Chin, W. S., Dekel Y. & Zhu, Y., Machine learning at microsoft with ml. net. *In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2448-2458, https://doi.org/10.1145/3292500.3330667 (2019).
6. Bogner K., Pappenberger F. & Zappa M. (2019). Machine learning techniques for predicting the energy consumption/production and its uncertainties driven by meteorological observations and forecasts. *Sustainability*, **11(12)**, 3328, doi:10.3390/su11123328 (2019).
7. Zahid, M., Chen, Y., Jamal, A. & Mamadou, C. Z.. Freeway short-term travel speed prediction based on data collection time-horizons: A fast forest quantile regression approach. *Sustainability*, **12(2)**, 646, doi:10.3390/su12020646 (2020).
8. Yang, J., Liu, Y., Liu, Z., Wu, Y., Li, T., & Yang, Y., A Framework for Human Activity Recognition Based on WiFi CSI Signal Enhancement. *International Journal of Antennas and Propagation,* 6654752 https://doi.org/10.1155/2021/6654752 (2021).
9. Hu, P., Yang, W., Wang, X., & Mao, S. (2019, December). MiFi: Device-Free Wheat Mildew Detection Using Off-the-Shelf WiFi Devices. *IEEE Global Communications Conference (GLOBECOM)*, 1-6, DOI: 10.1109/GLOBECOM38437.2019.9013895 (2019).
10. Deng, J., Sun, W., Guan, L., Zhao, N., Khan, M. B., Ren, A. &Abbasi, Q. H. Non-invasive suspicious liquid detection using wireless signals. Sensors, **19(19),** 4086, doi:doi.org/10.3390/s19194086 (2019).
11. Dang, X., Cao, Y., Hao, Z., & Liu, Y. WiGId: Indoor Group Identification with CSI-Based Random Forest. *Sensors*, **20(16)**, 4607, doi:10.3390/s20164607 (2020).
12. Ma, Y., Zhou, G., & Wang, S., WiFi sensing with channel state information: A survey. *ACM Computing Surveys (CSUR),* **52(3)**, 1-36 doi://doi.org/10.1145/3310194 (2019).
13. Song, H., Wei, B., Yu, Q., Xiao, X., &Kikkawa, T. WiEps: Measurement of Dielectric Property With Commodity WiFi Device—An Application to Ethanol/Water Mixture. *IEEE Internet of Things Journal*, **7(12)**, 11667-11677 doi:10.1109/JIOT.2020.2999210 (2020).
14. Li, F., Valero, M., Shahriar, H., Khan, R. A., &Ahamed, S. I., Wi-COVID: A COVID-19 symptom detection and patient monitoring framework using WiFi. *Smart Health*, **19**, 100147, doi:10.1016/j.smhl.2020.100147 (2021).
15. Atif, M., Muralidharan, S., Ko, H., &Yoo, B., Wi-ESP—A tool for CSI-based Device-Free Wi-Fi Sensing (DFWS). *Journal of Computational Design and Engineering*, **7(5)**, 644-656, doi:https://doi.org/10.1093/jcde/qwaa048 (2020).
16. Hernandez, S. M. & Bulut, E., Performing WiFi sensing with off-the-shelf smartphones. *IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops*), 1-3 **doi:** 10.1109/PerComWorkshops48775.2020.9156194 (2020).
17. Jadhav M. M. "Machine learning based autonomous fire combat turret." Turkish Journal of Computer and Mathematics Education (TURCOMAT), **12**, 2372-2381, https://doi.org/10.17762/turcomat.v12i2.2025 (2021)
18. Islam, Minarul, Md Shawkat Ali, Nusrat Jahan Shoumy, Sabira Khatun, Mohamad Shaiful Abdul Karim, and Bifta Sama Bari. "Non-invasive blood glucose concentration level estimation accuracy using ultra-wide band and artificial intelligence." SN Applied Sciences 2, 1-9, DOI:10.1007/s42452-019-1884-3 (2020)
19. Baghelani, Masoud, Zahra Abbasi, Mojgan Daneshmand, and Peter E. Light. "Non-invasive continuous-time glucose monitoring system using a chipless printable sensor based on split ring microwave resonators." *Scientific Reports* 10, no. 1, 12980, DOI:10.1038/s41598-020-69547-1 (2020):
20. Jain, P., Joshi, A. M., Agrawal, N., & Mohanty, S. (2020). iGLU 2.0: A new non-invasive, accurate serum glucometer for smart healthcare. arXiv preprint arXiv,09182, doi: https://doi.org/10.48550/arXiv. (2001).
21. Islam, M., Min, C. L., Shoumy, N. J., Ali, M. S., Khatun, S., Karim, M. S. A., & Bari, B. S. Software module development for non-invasive blood glucose measurement using an ultra-wide band and machine learning. In Journal of Physics: Conference Series. IOP Publishing, Vol. 1529, No. 5, 052066 9, DOI 10.1088/1742-6596/1529/5/052066 (2020).
22. Zahid, M., Chen, Y., Jamal, A., & Mamadou, C. Z.. Freeway short-term travel speed prediction based on data collection time-horizons: A fast forest quantile regression approach. Sustainability, 12(2), 646, https://doi.org/10.3390/su12020646 (2020).
23. Yang, J., Liu, Y., Liu, Z., Wu, Y., Li, T., & Yang, Y.. A Framework for Human Activity Recognition Based on WiFi CSI Signal Enhancement. International Journal of Antennas and Propagation, 6654752, https://doi.org/10.1155/2021/6654752 (2021).
24. Dang, X., Cao, Y., Hao, Z., & Liu, Y. WiGId: Indoor Group Identification with CSI-Based Random Forest. Sensors, 20(16), 4607, doi: 10.3390/s20164607 (2020).
25. Song, H., Wei, B., Yu, Q., Xiao, X., & Kikkawa, T. WiEps: Measurement of Dielectric Property With Commodity WiFi Device—An Application to Ethanol/Water Mixture. IEEE Internet of Things Journal, 7(12), 11667-11677, DOI: 10.1109/JIOT.2020.2999210 (2020).
26. Li, F., Valero, M., Shahriar, H., Khan, R. A., & Ahamed, S. I. Wi-COVID: A COVID-19 symptom detection and patient monitoring framework using WiFi. Smart Health, 19, 100147, DOI: 10.1016/j.smhl.2020.100147 (2021).
27. Atif, M., Muralidharan, S., Ko, H., & Yoo, B. Wi-ESP—A tool for CSI-based Device-Free Wi-Fi Sensing (DFWS). Journal of Computational Design and Engineering, 7(5), 644-656, https://doi.org/10.1093/jcde/qwaa048 (2020).
28. Reddy, Shiva Shankar, Nilambar Sethi, R. Rajender, and V. Sivarama Raju Vetukuri. "Non-invasive Diagnosis of Diabetes Using Chaotic Features and

Genetic Learning." In Third International Conference on Image Processing and Capsule Networks: ICIPCN 2022, pp. 161-170, DOI:10.1007/978-3-031-12413-6_13 (2022).

29. Agrawal, Harshita, Prateek Jain, and Amit M. Joshi. "Machine learning models for non-invasive glucose measurement: towards diabetes management in smart healthcare." Health and Technology 12, no. 5, 955-970, DOI: 10.1007/s12553-022-00690-7 (2022).

30. Bavkar, Vandana C., and Arundhati A. Shinde. "Machine learning algorithms for diabetes prediction and neural network method for blood glucose measurement." Indian J Sci Technol 14, no. 10, 869-880, DOI: 10.17485/IJST/v14i10.2187 (2021)