Heliyon 11 (2025) e41199

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

# Evaluation of machine learning-based regression techniques for prediction of diabetes levels fluctuations

Badriah Alkalifah [a], Muhammad Tariq Shaheen [b], Johrah Alotibi [a], Tahani Alsubait [a,*], Hosam Alhakami [c]

[a] Department of Computer Science and Artificial Intelligence, College of Computing, Umm Al-Qura University, Makkah, Saudi Arabia
[b] School of Electrical Engineering and Computer Science, NUST, Islamabad, Pakistan
[c] Department of Software Engineering, College of Computing, Umm Al-Qura University, Makkah, Saudi Arabia

## ARTICLE INFO

Dataset link: https://github.com/Almutaani/ML_Diabetes-.git

*Keywords:*
Diabetic
Continuous glucose monitoring (CGM)
Hypoglycemia
Hyperglycemia
Regression
Machine learning (ML)

## ABSTRACT

Middle-Aged and Elderly people today face a variety of health problems as a result of their modern lifestyle, which includes increased work stress, less physical activity, and altered food habits. Because of Complications arising, diabetes has become one of the most frequent, severe, and fatal illnesses around the world. Therefore, inaccurate measurements of blood glucose levels can seriously damage vital organs. Several strategies for long-term glucose prediction have been proposed in the literature. Unfortunately, these methods require the patient to identify their daily activities, which can be error-prone, such as meal intake, insulin injection, and emotional aspects. This paper suggests using continuous glucose monitoring (CGM) of 14733 patients, with three assistance factors to predict blood glucose levels independently of other parameters, hence reducing the burden on the patients. To support this an Artificial Neural Network (ANN), Binary Decision Tree (BDT), Linear Regression (LR), Boosting Regression Tree Ensemble (BSTE), Linear Regression with Stochastic Gradient Descent (LRSGD), Stepwise (SW), Support Vector Machine (SVM), and Gaussian process regression (GPR) were investigated. The result indicated that The highest classification accuracy of (92.58%) has been achieved by BDT followed by BSTE (92.04%) and GPR (88.59%). The obtained average of root means square error (MSE) was 1.64, 1.67, 1.69, mg/dL for prediction horizon (PH) respectively to GPR, BSTE, and ANN.

## 1. Introduction

As shown in a report published in 2017 by the International Diabetes Federation [1], there have been 537 million diabetics around the globe, with predictions that this number would rise to 783 million by the year 2045 [2]. Diabetes also contributes to ten percent of worldwide health spending (USD 760 billion) [3]. Diabetes Mellitus (DM) is a life-threatening, incurable health condition instigated by the inability of pancreas' beta cells (cells) to produce enough functional insulin for the body to utilize, resulting in chronically high blood glucose levels [4]. The insulin-glucose-intake cycle is depicted in Fig. 1. There is a chronic course to the disease, as well as a disturbance of all types of metabolism.

* Corresponding author.
*E-mail address:* tmsubait@uqu.edu.sa (T. Alsubait).

https://doi.org/10.1016/j.heliyon.2024.e41199
Received 17 March 2024; Received in revised form 6 December 2024; Accepted 12 December 2024

Available online 16 December 2024
2405-8440/

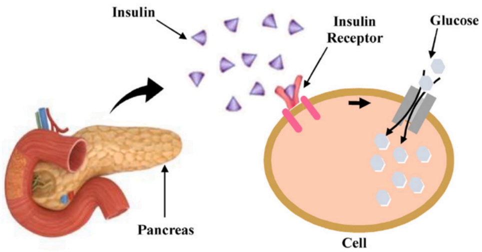


**Fig. 1.** The cycle of insulin on glucose intake [5].

The development of DM in patients has been linked to a number of variables. The current global increase in diabetes has been attributed to indolent routines, chubbiness, an absence of knowledge, and other variables [6]. There are three kinds of diabetes: gestational diabetes, diabetes mellitus of type-2 (T2D), and, diabetes mellitus of type-1 (T1D) [7,8].

Hypoglycemia occurs when the body generates too much insulin, causing low blood sugar concentrations. It is more common in patients with type-1 diabetes, but it can also happen in individuals with diabetes type who are on insulin or other medications. Low blood sugar can cause shakiness, hunger, a fast heart rate (tachycardia), and perspiration. Hypoglycemia symptoms occur progressively and may go overlooked at first, but they will become more deadly if left untreated.

Hyperglycemia occurs when glucose levels in the blood are elevated. It happens when the body doesn't have sufficient insulin or can't employ the insulin it has to convert glucose to energy. Hyperglycemia can cause fatigue, headaches, excessive urination, and extreme thirst. Over time, symptoms such as vomiting and diarrhea, breathlessness, and unconsciousness may appear.

Many patients are fearful of experiencing hypoglycemia (low blood sugar) and hyperglycemia (high blood sugar) which are categorized as: normal glycemia (the level of glucose is $\geq$ 70 mg/dL and <250 mg/dL), Hypoglycemia (the level of glucose is < 70 mg/dL), and Hyperglycemia (blood glucose is $\geq$ 250 mg/dL) [9]. More ways to diminish the hazard of hypoglycemia in diabetic patients have emerged as a result of recent advances in glucose monitoring and medication development [10]. As a result, for all people who have diabetes, Hypoglycemia and Hyperglycemia prediction [11] and early classification [12] are essential.

Currently, current cutting-edge insulin delivery methods combine glycemic data from continuous glucose monitors (CGMs), that assess levels of glucose in the subcutaneous interstitial fluid, along with pumps of insulin that give short-acting insulin formulations. The latter technology's use is contingent on the patient's effort of data on diet quantities and carbs, insulin bolus dose, exercise, and so on [13]. In the recent decade, several statistical approaches for forecasting future glucose readings from CGM data have indeed been proposed, including dynamic hazard [14], stochastic models [15,16] auto-regressive models [17–19] dynamic state-space models, and control algorithms [20,21], neural network-based models [22,23], support vector regression models [24,25] and machine learning (ML)-based models [19,25,26].

Hypoglycemia and hyperglycemia must be predicted accurately using the previous records to receive proper treatment and prevent the condition from escalating to severe consequences. The goal of this research is to develop models that are based on machine learning that can precisely predict the likelihood of low blood sugar "hypoglycemia" and high blood sugar "hyperglycemia" based on sugar readings from the previous days. It also concentrates on predicting the protracted hazard of iatrogenic hypoglycemia and hyperglycemia using data from continuous glucose monitors.

The structure of the paper is organized as follows. A brief summary of published CGM prediction methods is presented in Section 2. The Methodology for predicting blood glucose levels is described in Section 3. Section 4 presents the Experimental setup. The result is discussed in section 5. Finally, the conclusions, Limitations, and future works are provided in Section 6, 7 and 8 respectively.

## 2. Glucose prediction algorithms: state of the art

One of the most essential areas for society to progress via scientific knowledge [27,28] and technologies [29–31] is health care. Machine learning (ML) and Deep learning (DL) approaches [32–35] are strong tools that augment classical machine learning by allowing computers to learn through data [36] to develop smarter apps, process health records and apply computer vision for medical images and genomics [37].

There have been numerous attempts to construct models for forecasting blood glucose levels in diabetics. Natural online use of CGM is the prevention of hypo/hyperglycemic episodes. Some projection methods were proposed a few years following the introduction of CGM sensors to create alerts when the actual trend of the glucose concentration profile revealed that hypoglycemia was likely to occur within a short period. As an example, for such a scope, Choleau et al. [38] used a first-order linear extrapolation of the latest two/three glucose samples. A more complex strategy is to generate hypo/hyper warnings based on a glucose concentration forecast generated ahead of time from previous CGM data. Insulin pumps without and with integrated CGMs have been recommended to lower the risk of hypoglycemia, according to Johnson-Rabbett, and Seaquist [39]. In a three-month study, 247 experienced T1D patients of the pump were randomly allocated to a sensor-enhanced pump without or with low glucose suspend option, with the threshold suspend group having a 37.5% reduced mean area under the curve (AUC) for nocturnal hypoglycemia incidents. Nestoras et al. [40]. Researchers developed a model to predict the probability of iatrogenic hypoglycemia within 24 hours (the short-term risk) following each measurement of blood glucose (BG) throughout hospitalization. A data set of 54978 admissions of 35147 T1D patients from the biggest hospital was obtained and divided into a training set of 70% and a test set of 30%. Random Forest Classification (RF),

**Table 1**
Memorization of previous works.

| N0. | Dataset | Method | Features | Evaluation Results |
|---|---|---|---|---|
| [39] | 14 T1D. | AUC. | 3 | MAE = 37.5 mg/dl. |
| [40] | 35147 T1D. | RF, LR, naïve Bayes, and SGB. | 9 | ACC = 95% for (RF) model. |
| [41] | 22 T1D. | AR, and ARMA. | 4 | ACC = 90% . |
| [42] | 10 T1D. | NN. | 4 | RMSE = 44 mg/dl, MAPE = 22 |
| [9] | 11 T1D. | SML - based algorithm | 4 | RMSE = 21 mg/dl, Acc= 93%. |
| [43] | 130 US hospitals. | LR, BDT, SVM, NN, and DF. | 13 | ACC = 76% for (BDT) model. |
| [44] | 6,579 T1D. | Xg Boost, RF, NN, and CRT. | 14 | RMSE = (5.4 mg/dl) for (Xg Boost) model. |
| [5] | 12 T1D. | SVM. | 1 | RMSE = 45 mg/d. |
| [45] | 28 T1D. | AR1 | 4 | MSPE = 3.79 mg/d. |
| [46] | 20 virtual and 9 real patients. | NNLPA, NNPG, LR. | 3 | Real data: RMSE = 9.4 mg/dl and Virtual data: RMSE = 9.7 mg/dl, for (NNLPA) model. |

multivariable logistic regression, naive Bayes, stochastic gradient boosting were all examined for their effectiveness in predicting the outcome. The results of the internal validation were better than those of the external validation. The model had a C statistic of 0.90 (95 percent confidence interval: 0.89–0.90), a positive predictive value of 0.09 (95 percent confidence interval: 0.08–0.09).

Based on data acquired from monitoring 22 T1D patients over 48 hours, Eren-Oruklu et al. [41] built a model employing autoregressive moving average (ARMA) models. The purpose of this work is to advance consistent subject-based glucose forecast models employing CGM data. The accuracy of the predictions is measured using the glucose forecast error and the Clarke Error Grid analysis (CG-EGA). The CG-EGA analysis produces 90% incorrect values. The quality and quantity of the training dataset for networks have a significant impact on the ANN prediction results.

To address this issue, Pappada et al. [42] introduced a real-time glucose prediction feed-forward neural network model (NNM) consisting of a feed-forward network of (9 hidden neurons) with tangent (sigmoid) activation functions and (1 output) with a linear transfer function. The model was tested on 10 T1D patients who were being monitored by the CGMS Gold (Medtronic, Northridge, CA), the NNM was trained using a collection of 23,432 vectors of CGM and computerized diary data. MSE 22.1 mg/dl and RMSE 43.9 mg/dl were different percentages in NNM. The work of Yonit et al. [9]. SML analysis, based on a continuous glucose tracking system, boosted blood glucose concentration forecasting and future hypoglycemia and hyperglycemia occurrences (CGM). The data was acquired from 11 T1D individuals between the ages of 18 and 39 during a period of 7 to 50 days. The algorithm's tailored solutions were successful in anticipating glucose concentration 30 minutes from the last check.

The true-positive-hypoglycemia-prediction-rate was 64 percent using the best-fit model, while the false-positive-rate was 4.0 percent and the false-negative-rate was 0.015 percent. The real rate of positive hyperglycemia prediction was 61%. SML tools that are state-of-the-art are successful in forecasting the glucose levels of people with type-1 diabetes and alerting them to future hypoglycemia and hyperglycemia episodes, thereby improving glycemic management. In Reid [43]. The dataset spans ten years (1999–2008) and includes 50 attributes and 101,766 occurrences from 130 US institutions. LR, BDT, SVM, NN, and DF are examples of predictive models. The performance of the models where a BDT got the best results was measured using metrics such as accuracy, recall, and AUC. The accuracy of hospital readmission is 86%, and the accuracy of diabetes diagnosis is 67%. To predict and improve performance after surgery, Akira et al. [44] built and evaluated machine learning and linear regression approaches, random Forest (RF), Extreme Gradient Boosting (Xg Boost), neural network (NN), regression trees and Classification. A substantial perioperative dataset (6,579 T1D patients) comprising patient- and surgery-specific information was utilized to train the model, with 5,265 (80%) being used for training and 1,314 (20%) being used for validation. The model with the best results, Xg Boost, was turned into a web tool called Hyper-G, which perioperative practitioners may use to anticipate peak glucose concentration during operations at the local clinic.

To develop a more reliable prediction compared to the cited works, a recent approach based on support vector regression (SVR) was proposed in Hamdi et al. [5] combining support vector regression with a differential evolution technique in a model The researcher recommends utilizing solely continuous glucose monitoring (CGM) data to predict blood glucose levels, regardless of other factors. The recommended technique is validated using real CGM data from 12 T1D patients. The root mean square error (RMSE) for prediction horizons (PH) 45 mg/dl.

Sparacino et al. (2010) [45] studied an AR-based model in a hospital setting. For 48 hours, 28 T1D participants were monitored every three minutes by the Glucoday, a noninvasive CGM device, to test if they could predict glucose concentration ahead of time. With a time window of 30 minutes, the results demonstrate that breaching the hypoglycemic threshold may be predicted 20–25 minutes in advance. With the AR model, the performance of the algorithms was tested using both classic signals estimating indices and novel delay indices, and the MSPE was 3.79 mg/d.

Zecchin et al. (2012) [46] developed a classifier using a neural network model combined with a linear prediction algorithm (NN-LPA) and a first-order autoregressive model (NNPG). This model integrates historical CGM data with carbohydrate intake information. The prediction algorithm was tested on 20 virtual patients created using a type-1 diabetes simulator and on 9 real datasets from type-1 diabetics monitored for 7 days with the Abbott FreeStyle NavigatorTM CGM sensor. The best RMSE for real data was 9.4 mg/dl over 5 minutes, while for virtual data it was 9.7 mg/dl over 5 minutes. Table 1 summarizes the studies on predicting blood glucose levels using ML techniques.

Recent advancements in machine learning and data analysis have led to significant progress across various domains, enhancing our understanding and treatment of complex health issues. For instance, optimizing epileptic seizure recognition performance through feature scaling and dropout layers has shown promise in improving predictive accuracy [47]. Similarly, the acoustic detection of large-

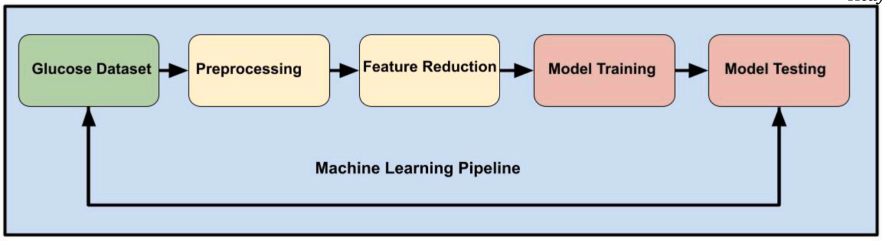


**Fig. 2.** The utilized machine learning pipeline to process the blood glucose data.

scale emergency vehicle sirens amidst road noise has highlighted the potential of machine learning in public safety applications [48]. In women's health, revolutionizing core muscle analysis related to female sexual dysfunction through machine learning offers new insights into effective therapeutic strategies [49]. Additionally, utilizing convolutional neural networks to classify monkeypox skin lesions represents a critical step in dermatological diagnosis [50]. On the oncology front, harnessing machine learning to identify synergistic combinations of FDA-approved cancer drugs could pave the way for more effective treatment protocols (Singh et al., 2023). Furthermore, optimizing disease classification through language model analysis of symptoms provides a novel approach to enhancing diagnostic precision [51]. Research on predicting female pelvic tilt and lumbar angle in cases of urinary incontinence and sexual dysfunction underscores the intersection of biomechanics and machine learning [52]. Lastly, a machine learning framework for predicting Hepatitis C Virus outcomes has yielded valuable insights in a real-world case study conducted in Egypt, demonstrating the practical applications of these technologies in global health [53].

## 3. Methodology

Fig. 2 depicts the methodological framework used in this study, which comprises five major stages: data collection, pre-processing, feature reduction, model training, and model testing for blood glucose level prediction.

In the first stage, data collection, raw datasets are acquired. During the second stage, pre-processing, occurrences with non-available (NA) values are removed. The third stage, feature reduction, involves eliminating redundant variables to prevent overfitting of the machine learning model. The fourth stage, model training, involves training the machine learning model with the prepared dataset. Finally, in the fifth stage, the trained model is tested to predict blood glucose levels.

### *3.1. Dataset*

The dataset utilized in this study comprises records of individuals' blood glucose levels alongside their superficial body feature readings [54]. It includes 16,969 records of individuals of various ages, among which 16,641 are diabetic and 328 are non-diabetic. The dataset encompasses superficial bodily parameters such as body temperature, heart rate, blood pressure, pulse oximetry (SPO2), sweating, and shivering. The primary objective of this dataset is to understand the impact of blood glucose levels on these superficial body markers. Variations in blood glucose levels significantly influence these physiological parameters. The dataset is a combination of authentic and synthetic recordings [55]. True data were collected from diabetic patients using electronic devices to measure body temperature, heart rate, and blood pressure. Additionally, real records of diabetic patients were obtained through finger pricking or using flash glucose monitoring tools. For non-diabetic individuals, data were collected using electronic devices such as smart wristbands. Their synthetic blood glucose levels were calculated based on the average blood glucose level over five days, which showed minimal variance in non-diabetic patients.

### *3.2. Data pre-processing*

The dataset underwent crucial pre-processing due to its collection from various monitoring devices. Initially comprising 16,969 records, it was refined and cleaned to 14,733 records, delineated into two categories: diabetic and non-diabetic

### *3.3. Feature reduction*

#### *3.3.1. Neighborhood component analysis*

Numerous studies have shown that individuals with different traits exhibit varying discrimination abilities. However, simply utilizing the entire array of collected characteristics during the training phase often results in subpar sensitivity when estimating blood glucose levels. Some features may even have a detrimental effect on the accuracy of predictions. Therefore, it's imperative to identify the key criteria that make a machine learning (ML) model effective. This typically involves investigating the suitability of each characteristic for the task of predicting blood glucose levels. Through methods like neighborhood component analysis (NCA), researchers can pinpoint the most relevant features that contribute to accurate glucose level estimations, as depicted in Fig. 3.

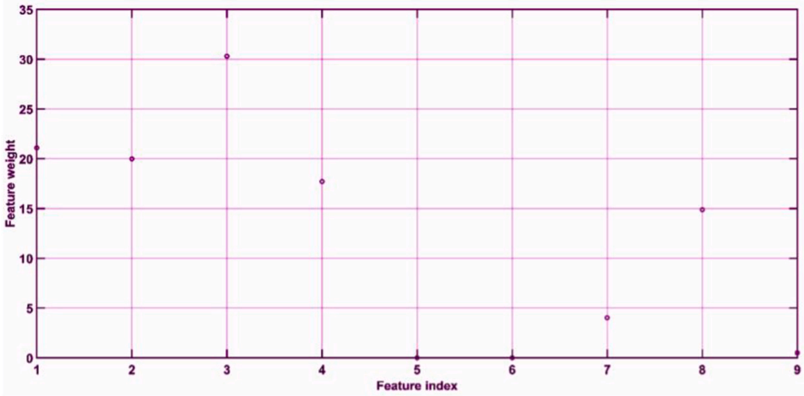


**Fig. 3.** The feature weight of each investigated feature to predict the blood glucose level.

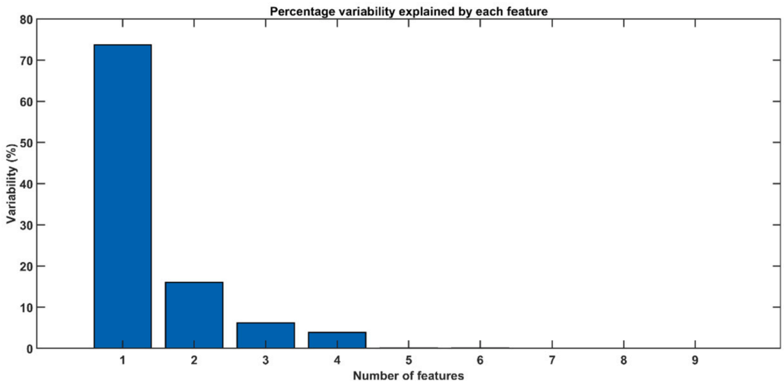


**Fig. 4.** The percentage variability of each investigated feature to predict the blood glucose levels.

### *3.3.2. Principle component analysis*

In an optimal feature selection process, the goal is to gather a small number of attributes with high variance and minimal correlation among traits. Principal Components Analysis (PCA), as described by Pino et al. (2021), determines the most effective collection of characteristics. It generates a set of key components that are uncorrelated and compact in size, ordered by variance rank (where the first component exhibits more variation than the second, and so forth). In this study, PCA was employed to streamline the information set from nine features (Blood Glucose Level (BGL), body temperature, heart rate, blood pressure, Pulse oximetry, Sweating, Shivering, Diabetic/Non-Diabetic status, and age) down to just four features (Blood Glucose Level (BGL), body temperature, heart rate, and blood pressure). This reduction aimed to decrease complexity and mitigate the risk of overfitting. The selected four elements encapsulate 99 percent of the dataset's variation. Fig. 4 illustrates the percentage variation of each of the nine criteria. It becomes evident that the first four attributes are closely associated with blood glucose readings, while the remaining attributes are deemed redundant.

## *3.4. Splitting data*

To impartially evaluate machine learning models, data is typically split into training and testing subsets. An empirical study suggests that allocating 30% of the data for validation and testing, and the remaining 70% for training, yields the most accurate predictions. In line with this recommendation, the dataset was randomly divided using the devset function: 70% for training (10,313 records), 15% for validation (2,010 records), and 15% for testing (2,010 records). This approach ensures that the model's performance is assessed on data it hasn't been exposed to during training, guaranteeing fairness in the analysis.

## *3.5. Regression analysis*

Various regression models were employed to predict blood glucose levels, including Gaussian process regression (GPR), linear regression with stochastic gradient descent (LRSGD), boosting regression tree ensemble (BSTE), binary decision tree (BDT), linear regression (LR), stepwise (SW), artificial neural network (ANN), and support vector machine (SVM). The proposed artificial neural network (ANN) model, illustrated in Fig. 5, utilized a feed-forward architecture. It underwent training using the Levenberg-Marquardt training method and the log transfer function. This model comprises one input layer with three neurons, a single hidden layer with ten neurons employing tangent (ReLU) activation functions, and one output layer with a linear transfer function. This research presents

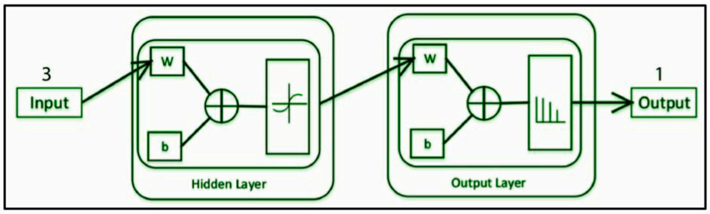


**Fig. 5.** The structure of the neural network in MATLAB.

a real-time glucose prediction feed-forward neural network model (NNM), designed specifically for accurate blood glucose level prediction.

The paper presents a feed-forward neural network (ANN) designed for predicting blood glucose levels. It features a simple architecture with an input layer (3 neurons), a single hidden layer (10 neurons) using ReLU activation for feature extraction, and an output layer (1 neuron) with linear activation for continuous prediction. The Levenberg-Marquardt training method is employed, supported by a log transfer function for preprocessing. This design emphasizes simplicity, effectiveness in handling nonlinear relationships, and suitability for real-time glucose prediction, reflecting a balance between model complexity and computational efficiency tailored to the task at hand.

## 4. Experimental setup

The experiments were conducted using MATLAB 2020, a software platform developed by MathWorks Inc., renowned for its capabilities in algorithm development, data visualization, and analysis. Leveraging an advanced computational language and an interactive numerical computation environment, MATLAB is widely utilized in scientific research. The hardware specifications of the computing device utilized for the experiments include an iMac (Retina 4K, 21.5-inch, 2017) equipped with a 3.4 GHz Quad-Core Intel Core i5 processor and 8 GB 2400 MHz DDR4 memory. A training dataset was employed to train the machine learning models, while a separate testing dataset was used to evaluate their performance. The dataset was randomly partitioned into training (70%), validation (15%), and testing (15%) subsets using the "devset" function. Principle component analysis (PCA) was implemented to mitigate overfitting by artificially oversampling the training samples. The "quantum particle swarm (QPS)" algorithm served as the optimizer during model training. The training process involved an elapsed time of 00:00:11, spanning 90 epochs with six validation checks. The overall hyperparameters for the used models are given in the following table.

### *4.1. Performance evaluation*

For the experiments evaluated, an accuracy matrix was used to evaluate the models' performance in addition to the confusion matrix. The Equation (1) shows the formula of accuracy:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \tag{1}$$

Where TP, FN, TN, and FP represent the true positive prediction, false negative prediction, true negative prediction, and false positive prediction, respectively.

Several quality measures are offered to examine the effectiveness of the regressor to analyze the effectiveness of regression investigation. For both learning and evaluation experiments, mean absolute value (MAE) shown in Equation (4), mean squared error (MSE) represented in Equation (3), and root mean squared error (RMSE) mathematically represented in Equation (2) were determined to assess and compare the efficacy of individual regressors.

$$RMSE = \sqrt{\frac{1}{n}\sum_{n=1}^{t=1} e_t^2} \tag{2}$$

$$MSE = \frac{1}{n}\sum_{n=1}^{t=1} e_t^2 \tag{3}$$

$$MAE = \frac{1}{n}\sum_{n=1}^{t=1} |e_t| \tag{4}$$

## 5. Results

### *5.1. Blood glucose level predictions*

Table 2 presents the efficacy of blood glucose concentration projections using different training strategies, while Table 3 evaluates various machine learning networks. Among these approaches, GPR demonstrates the highest efficacy with an MSE of 1.64 mg/dl,

**Table 2**
Dataset description for blood glucose and physiological parameters study.

| Feature | Description |
|---|---|
| **Total Records** | 16,969 individual records |
| **Diabetic Records** | 16,641 records of diabetic individuals |
| **Non-Diabetic Records** | 328 records of non-diabetic individuals |
| **Age Range** | Various ages |
| **Blood Glucose Levels** | Collected through finger pricking or flash glucose monitoring for diabetic individuals; synthetic values for non-diabetics |
| **Body Temperature** | Measured using electronic devices |
| **Heart Rate** | Recorded via electronic monitoring |
| **Blood Pressure** | Monitored with electronic devices |
| **Pulse Oximetry (SPO2)** | Collected via electronic measurement |
| **Sweating** | Recorded as a physiological marker |
| **Shivering** | Monitored as a physiological response |
| **Data Composition** | Authentic data for diabetic records, synthetic for non-diabetic blood glucose levels |
| **Non-Diabetic Data Source** | Collected with smart wristbands; synthetic glucose based on a five-day average |

**Table 3**
Hyperparameters of different models.

| Model_Algorithm | Hyperparameters used | Values_Description |
|---|---|---|
| Neighborhood Component Analysis (NCA) | Lambda | 0.5 |
| Neural_Network (NN) | Training_function | trainscg |
| | Hidden_layer_size | 1 |
| Stochastic_Gradient_Descent | Learner | leastsquares |
| Gaussian_Process_Regression | Basis_function | constant |
| | Kernal_function | exponential |
| Ensemble_of_Learners_(Boosted_trees) | No_of_learning_cycles | 30 |
| | Learning_rate | 0.1 |
| | Minimum_leaf_size | 8 |
| Decision_tree (BDT) | Minimum_leaf_size | 4 |
| Step_wise_regression | No_of_steps | 1000 |

followed by BSTE with an MSE of 1.67 mg/dl, and ANN with an MSE of 1.69 mg/dl. Lower MAE, MSE, and RMSE values indicate better performance of the regression methods.

The reliability statistics of all explored regression approaches on testing data are depicted in Fig. 6, blood glucose concentrations, both actual and anticipated, are graphically shown. The performance of LR, LRSGD, GPR, BSTE, SW, and ANN may be seen to be equivalent. The SVM and BDT, on the other hand, have the poorest productivity outcomes.

### *5.2. Hypoglycemia and hyperglycemia predictions*

Along with forecasting blood glucose levels, the study's secondary goal was to categorize the glucose levels into the appropriate glucose classifications, such as hyperglycemia, hypoglycemia, and normal. Hypoglycemia (blood glucose less than 70 mg/dL), normoglycemia (blood glucose between 70 mg/dL and 250 mg/dL), and hyperglycemia (blood glucose greater than 250 mg/dL). In order to determine which machine learning model predicts the proper class of blood glucose level based on the dataset.

Fig. 7 shows the outcomes of the tests in the form of confusion matrices. Used the classification accuracy of the performance measuring parameter in this study to evaluate the performance of the machine learning algorithms. From Fig. 7. The highest classification accuracy observed (92.58%) has been achieved by BDT followed by BSTE (92.04%) and GPR (88.59%). Also, it should be observed that since the utilized dataset is imbalanced concerning the number of classes (hypoglycemia, hyperglycemia, and normal), the poor performance of the machine learning models is due to the inclusion of the hyperglycemia class. Thus it is suggested that in the future a balanced dataset should be explored to further investigate the hypothesis of the presented study. The table below presents the $R^2$ metric for regression models evaluated during both training and testing phases.

Table 4 shows the classification reliability of all machine learning strategies used for hypoglycemia, hyperglycemia, and normal class forecast. The greatest classification accuracy was attained by BDT, followed by BSTE, then GPR. It's worth mentioning that both BDT and BSTE are decision tree-based machine learning algorithms. The goal of using a Decision Tree is to construct a training framework that can be used to identify the class or amount of the attribute value by learning fundamental decision rules determined from prior data.

### *5.3. Comparison with previous literature on the same machine learning models*

It is a well-known and widely accepted rule of thumb to always compare the results of machine learning models within the same domain with the same and/or similar dataset and the same models. Even through this, it is the first time the same database has been used to predict and anticipate continuous blood glucose concentrations (in this study). As a result, the study's findings cannot

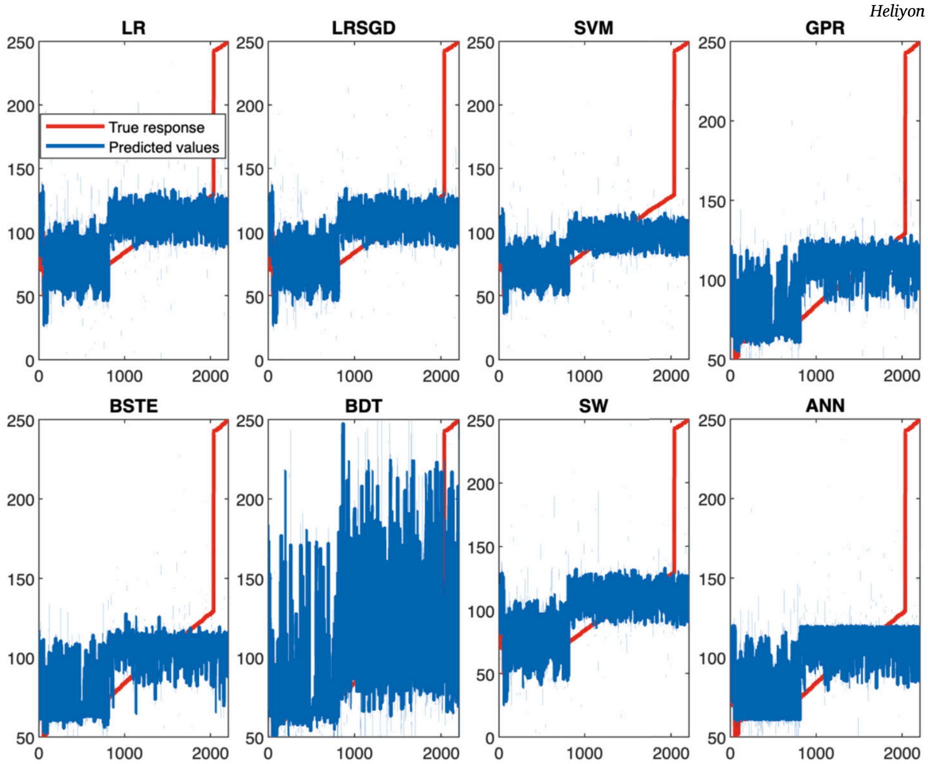


**Fig. 6.** Performance results of all investigated regression techniques for testing data. The blue lines represent actual blood glucose levels and the red lines represent predicted blood glucose levels.

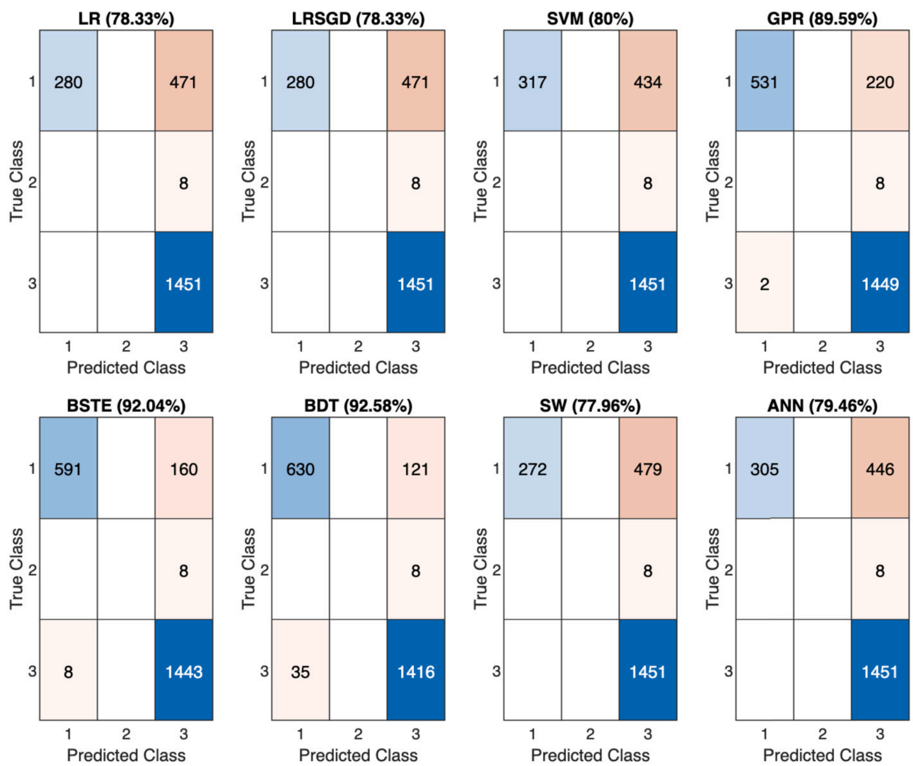


**Fig. 7.** Performance results of all investigated regression techniques for testing data. Hypoglycemia, hyperglycemia, and normal are denoted by the numbers 1, 2, and 3, respectively.

**Table 4**
ML regression techniques (training results).

| | MAE | MSE | RMSE |
|---|---|---|---|
| LR | 24.9587 | 1.6733e+03 | 40.92.0458 |
| LRSGD | 24.9584 | 1.6733e+03 | 40.92.0458 |
| SVM | 23.1019 | 1.8038e+03 | 42.4712 |
| **GPR** | 22.4098 | **1.5409e+03** | 39.2538 |
| **BSTE** | 21.6135 | **1.5727+03** | 39.6576 |
| BDT | 16.6680 | 872.5468 | 29.5389 |
| SW | 24.9671 | 1.6709e+03 | 40.8771 |
| **ANN** | 23.8636 | **1.6344e03** | 40.4281 |

**Table 5**
ML regression techniques (testing results).

| | MAE | MSE | RMSE |
|---|---|---|---|
| LR | 25.4673 | 1.7411e+03 | 41.7266 |
| LRSGD | 25.4670 | 1.7411e+03 | 41.7266 |
| SVM | 23.6007 | 1.8809e+03 | 43.3692 |
| **GPR** | 22.9470 | **1.6487e+03** | 40.6047 |
| **BSTE** | 22.0655 | **1.6706e+03** | 40.8728 |
| BDT | 28.1651 | 2.2681e+03 | 47.6241 |
| SW | 25.4976 | 1.7411e+03 | 41.7261 |
| **ANN** | 24.1311 | **1.692.049e03** | 41.1203 |

**Table 6**
Comparison of $R^2$ values for training and test sets in regression.

| Model | R2_Train | R2_Test |
|---|---|---|
| LinearReg | 0.1933 | 0.1905 |
| LinearSGD | 0.1933 | 0.1905 |
| SVM | 0.1315 | 0.1115 |
| GP | 0.2565 | 0.2265 |
| Ensemble | 0.2460 | 0.2113 |
| Tree | 0.5792 | -0.0439 |
| StepWise | 0.1944 | 0.1912 |
| NN | 0.2382 | 0.2113 |

**Table 7**
Prediction of hypo/hyperglycemia.

| Model | Accuracy |
|---|---|
| LR | 78.33% |
| LRSGD | 78.33% |
| SVM | 80% |
| **GPR** | **89.59** |
| **BSTE** | **92.04%** |
| **BDT** | **92.58%** |
| SW | 77.96% |
| ANN | 79.46% |

be linked to those of other research using the same dataset. To compare with previous work, we used the same models only and explained the performance evaluation. The results are shown in Table 5.

To compare with previous work according to classification for prediction of the three classes of hypoglycemia, normoglycemia, and hyperglycemia, Table 6 shows the result.

Based on all the classifier's results, it reveals the proposed BDT performs best over all other models with an accuracy of prediction the accurate results is 92.58%. Table 7 shows the overall accuracy of all the algorithms while Table 8 shows the comparison to previous work based on performance evaluation for prediction results. Moreover, Table 9 shows comparison of previous work for classification.

## 6. Conclusion

Clinicians must work hard to keep hypoglycemia and hyperglycemia under control while maintaining optimum management. Hypoglycemia and hyperglycemia must be precisely anticipated using records to obtain adequate treatment and prevent the disease

**Table 8**
Comparison to previous work based on performance evaluation for prediction results.

| Resarch | Model | Evaluation Metrics | | |
|---|---|---|---|---|
| | | MAE | MSE | RMSE |
| Hamdi et al. | SVM | N/A | N/A | 45 |
| **Our Work** | | 23.6 | 1.8 | 43.4 |
| Pappada et al. | ANN | N/A | N/A | 43.9 |
| **Our Work** | | 24.1 | 1.69 | 41.1 |
| Zecchin et al. | LR | N/A | N/A | 23.8 |
| **Our Work** | | 25.46 | 1.74 | 41.72 |

**Table 9**
Comparison of previous work for classification.

| Research | Accuracy | Features | Model |
|---|---|---|---|
| Patil et al. [56] | 90% | 8 | K-means, Decision Trees |
| Chen et al. [57] | 90.04% | 8 | Decision_Tree |
| Komi et al. 2017 [51] | 89% | 8 | Artificial Neural Network |
| Reid, Clodagh [43] | 67% | 13 | BDT |
| **Proposed Work** | **92.58%** | 4 | BDT |

from worsening. Early monitoring of blood glucose levels, prediction, and diagnosis of varying diabetes levels using machine learning are critical for prompt treatment. The goal of the work was to see how well machine learning-based regression approaches could anticipate blood glucose readings. Depending on four features: blood glucose, blood pressure, heart rate, and body temperature, eight machine learning-based prediction approaches, including LR, LRSGD, SVM, GPR, BSET, BDT, SW, and ANN, were used. The findings revealed that for blood glucose values, GPR (MSE = 1.64 mg/dL) and BEST (MSE = 1.67 mg/dL) provided the greatest outcomes. While investigating the prediction of hypoglycemia, hyperglycemia, and normal classes, it was observed that the highest classification accuracy (92.58%) was achieved by BDT, followed by BEST (92.04%) and GPR (89.59%). This experiment is innovative in and of itself and will serve as a foundation for future research by the authors.

## 7. Limitations

This study has some limitations. The findings are based on a small dataset, and it is difficult to collect balanced datasets from patients. The vast amount of data being gathered by monitoring architectures also causes data handling and processing problems. Because the present research is based on limited inspection or experimental data, the applicability and generalization are limited.

## 8. Future work

In the future, the proposed dataset will be increased by including more patients to create a larger validation dataset. The GPR, BDT, and BEST prediction models will need to be clinically validated in preparation for possible real-world use. While this study employed well-known ML techniques, there will also be work to add more ML and DL techniques to detect and predict blood glucose levels to help doctors' decision-making. To address the problem of big data generated from CGM devices, various machine learning techniques must be explored to achieve the best results for prediction.

## CRediT authorship contribution statement

**Badriah Alkalifah:** Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis, Data curation, Conceptualization. **Muhammad Tariq Shaheen:** Writing – review & editing. **Johrah Alotibi:** Writing – review & editing, Visualization, Resources, Methodology, Formal analysis, Data curation. **Tahani Alsubait:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Conceptualization. **Hosam Alhakami:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Conceptualization.

## Funding

The authors extend their appreciation to the Deanship for Research & Innovation, Ministry of Education, Saudi Arabia for funding this research work through the project number: IFP22UQU4250089DSR200.

## Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Hosam Alhakami reports financial support was provided by Kingdom of Saudi Arabia Ministry of Education. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Dataset and code have been deposited at Click here to access the dataset and code.

## References

[1] M. East, N. Africa, Idf diabetes Atlas, Diabetes 20 (2017) 79.
[2] G. Li, S. Peng, C. Wang, J. Niu, Y. Yuan, An energy-efficient data collection scheme using denoising autoencoder in wireless sensor networks, Tsinghua Sci. Technol. 24 (1) (2018) 86–96.
[3] P.B. Jensen, L.J. Jensen, S. Brunak, Mining electronic health records: towards better research applications and clinical care, Nat. Rev. Genet. 13 (6) (2012) 395–405.
[4] K.G. Alberti, P.Z. Zimmet, Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus. Provisional report of a who consultation, Diabet. Med. 15 (7) (1998) 539–553.
[5] T. Hamdi, J.B. Ali, V. Di Costanzo, F. Fnaiech, E. Moreau, J.-M. Ginoux, Accurate prediction of continuous blood glucose based on support vector regression and differential evolution algorithm, Biocybern. Biomed. Eng. 38 (2) (2018) 362–372.
[6] K.C. Shih, A.S. Kwong, J.H. Wang, J.K. Wong, W.W. Ko, J.S. Lai, J.C. Chan, Diabetic retinopathy screening during the coronavirus disease 2019 pandemic, Eye 34 (7) (2020) 1246–1247.
[7] M. Asad, U. Qamar, A. Khan, R.U. Safdar, A systematic literature review of continuous blood glucose monitoring and suggesting the quantity of insulin or artificial pancreas (ap) for diabetic type 1 patients, in: Proceedings of the 2019 11th International Conference on Machine Learning and Computing, 2019, pp. 539–545.
[8] A.I. Alharbi, V. Gay, M.J. AlGhamdi, R. Alturki, H.J. Alyamani, Towards an application helping to minimize medication error rate, Mob. Inf. Syst. 2021 (2021) 1–7.
[9] Y. Marcus, R. Eldor, M. Yaron, S. Shaklai, M. Ish-Shalom, G. Shefer, N. Stern, N. Golan, A.Z. Dvir, O. Pele, et al., Improving blood glucose level predictability using machine learning, Diabetes/Metab. Res. Rev. 36 (8) (2020) e3348.
[10] H. Liu, H. Kou, C. Yan, L. Qi, Link prediction in paper citation network to construct paper correlation graph, EURASIP J. Wirel. Commun. Netw. 2019 (1) (2019) 1–12.
[11] R. Miotto, F. Wang, S. Wang, X. Jiang, J.T. Dudley, Deep learning for healthcare: review, opportunities and challenges, Brief. Bioinform. 19 (6) (2018) 1236–1246.
[12] W.S. McCulloch, W. Pitts, A logical calculus of the ideas immanent in nervous activity, Bull. Math. Biophys. 5 (4) (1943) 115–133.
[13] S.R. Colberg, R. Laan, E. Dassau, D. Kerr, Physical activity and type 1 diabetes: time for a rewire?, J. Diabetes Sci. Technol. 9 (3) (2015) 609–618.
[14] S. Guerra, G. Sparacino, A. Facchinetti, M. Schiavon, C.D. Man, C. Cobelli, A dynamic risk measure from continuous glucose monitoring data, Diabetes Technol. Ther. 13 (8) (2011) 843–852.
[15] P. Magni, R. Bellazzi, A stochastic model to assess the variability of blood glucose time series in diabetic patients self-monitoring, IEEE Trans. Biomed. Eng. 53 (6) (2006) 977–985.
[16] T. Briegel, V. Tresp, A nonlinear state space model for the blood glucose metabolism of a diabetic (ein nichtlineares zustandsraummodell für den blutglukosemetabolismus eines diabetikers), 2002.
[17] A. Gani, A.V. Gribok, Y. Lu, W.K. Ward, R.A. Vigersky, J. Reifman, Universal glucose models for predicting subcutaneous glucose concentration in humans, IEEE Trans. Inf. Technol. Biomed. 14 (1) (2009) 157–165.
[18] A. Gani, A.V. Gribok, S. Rajaraman, W.K. Ward, J. Reifman, Predicting subcutaneous glucose concentration in humans: data-driven glucose modeling, IEEE Trans. Biomed. Eng. 56 (2) (2008) 246–254.
[19] H.N. Mhaskar, S.V. Pereverzyev, M.D. Van der Walt, A deep learning approach to diabetic blood glucose prediction, Front. Appl. Math. Stat. 3 (2017) 14.
[20] S.M. Lynch, B.W. Bequette, Estimation-based model predictive control of blood glucose in type I diabetics: a simulation study, in: Proceedings of the IEEE 27th Annual Northeast Bioengineering Conference (Cat. No. 01CH37201), IEEE, 2001, pp. 79–80.
[21] K. Turksoy, L. Quinn, E. Littlejohn, A. Cinar, Multivariable adaptive identification and control for artificial pancreas systems, IEEE Trans. Biomed. Eng. 61 (3) (2013) 883–891.
[22] S. Ghimire, T. Celik, M. Gerdes, C.W. Omlin, Deep learning for blood glucose level prediction: how well do models generalize across different data sets?, PLoS ONE 19 (9) (2024).
[23] C. Pérez-Gandía, A. Facchinetti, G. Sparacino, C. Cobelli, E. Gómez, M. Rigla, A. de Leiva, M. Hernando, Artificial neural network algorithm for online glucose prediction from continuous glucose monitoring, Diabetes Technol. Ther. 12 (1) (2010) 81–88.
[24] R. Bunescu, N. Struble, C. Marling, J. Shubrook, F. Schwartz, Blood glucose level prediction using physiological models and support vector regression, in: 2013 12th International Conference on Machine Learning and Applications, vol. 1, IEEE, 2013, pp. 135–140.
[25] M.P. Reymann, E. Dorschky, B.H. Groh, C. Martindale, P. Blank, B.M. Eskofier, Blood glucose level prediction based on support vector regression using mobile platforms, in: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC, IEEE, 2016, pp. 2990–2993.
[26] K. Plis, R. Bunescu, C. Marling, J. Shubrook, F. Schwartz, A machine learning approach to predicting blood glucose levels for diabetes management, in: Workshops at the Twenty-Eighth AAAI Conference on Artificial Intelligence, 2014.
[27] H. Liu, H. Kou, C. Yan, L. Qi, Keywords-driven and popularity-aware paper recommendation based on undirected paper citation graph, Complexity 2020 (2020).
[28] J. Li, T. Cai, K. Deng, X. Wang, T. Sellis, F. Xia, Community-diversified influence maximization in social networks, Inf. Syst. 92 (2020) 101522.
[29] Y. Huang, Y. Chai, Y. Liu, J. Shen, Architecture of next-generation e-commerce platform, Tsinghua Sci. Technol. 24 (1) (2018) 18–29.
[30] Q. Liu, Y. Tian, J. Wu, T. Peng, G. Wang, Enabling verifiable and dynamic ranked search over outsourced data, IEEE Trans. Serv. Comput. (2019).
[31] K.H. Almotairi, Application of internet of things in healthcare domain, J. Umm Al-Qura Univ. Eng. Archit. 14 (1) (2023) 1–12.
[32] X. Xu, X. Zhang, H. Gao, Y. Xue, L. Qi, W. Dou, Become: blockchain-enabled computation offloading for iot in mobile edge computing, IEEE Trans. Ind. Inform. 16 (6) (2019) 4187–4195.
[33] Y. Zhang, K. Wang, Q. He, F. Chen, S. Deng, Z. Zheng, Y. Yang, Covering-based web service quality prediction via neighborhood-aware matrix factorization, IEEE Trans. Serv. Comput. (2019).
[34] Y. Zhang, C. Yin, Q. Wu, Q. He, H. Zhu, Location-aware deep collaborative filtering for service recommendation, IEEE Trans. Syst. Man Cybern. Syst. 51 (6) (2019) 3796–3807.
[35] M. Ibrahim, M. Wedyan, R. Alturki, M.A. Khan, A. Al-Jumaily, et al., Augmentation in healthcare: augmented biosignal using deep learning and tensor representation, J. Healthc. Eng. 2021 (2021).
[36] L. Qi, W. Dou, C. Hu, Y. Zhou, J. Yu, A context-aware service evaluation approach over big data for cloud applications, IEEE Trans. Cloud Comput. 8 (2) (2015) 338–348.
[37] C. Zhou, A. Li, A. Hou, Z. Zhang, Z. Zhang, P. Dai, F. Wang, Modeling methodology for early warning of chronic heart failure based on real medical big data, Expert Syst. Appl. 151 (2020) 113361.

[38] C. Choleau, P. Dokladal, J.-C. Klein, W.K. Ward, G.S. Wilson, G. Reach, Prevention of hypoglycemia using risk assessment with a continuous glucose monitoring system, Diabetes 51 (11) (2002) 3263–3273.
[39] B. Johnson-Rabbett, E.R. Seaquist, Hypoglycemia in diabetes: the dark side of diabetes treatment. A patient-centered review, J. Diabetes 11 (9) (2019) 711–718.
[40] N.N. Mathioudakis, M.S. Abusamaan, A.F. Shakarchi, S. Sokolinsky, S. Fayzullin, J. McGready, M. Zilbermint, S. Saria, S.H. Golden, Development and validation of a machine learning model to predict near-term risk of iatrogenic hypoglycemia in hospitalized patients, JAMA Netw. Open 4 (1) (2021) e2030913.
[41] M. Eren-Oruklu, A. Cinar, L. Quinn, D. Smith, Estimation of future glucose concentrations with subject-specific recursive linear models, Diabetes Technol. Ther. 11 (4) (2009) 243–253.
[42] S.M. Pappada, B.D. Cameron, P.M. Rosman, R.E. Bourey, T.J. Papadimos, W. Olorunto, M.J. Borst, Neural network-based real-time prediction of glucose in patients with insulin-dependent diabetes, Diabetes Technol. Ther. 13 (2) (2011) 135–141.
[43] C. Reid, Diabetes diagnosis and readmission risks predictive modelling: USA, Ph.D. thesis, National College of Ireland, Dublin, 2019.
[44] A.A. Nair, M. Velagapudi, L. Behara, R. Venigandla, C.T. Fong, M. Horibe, B.G. Nair, Hyper-g: an artificial intelligence tool for optimal decision-making and management of blood glucose levels in surgery patients, Methods Inf. Med. 58 (02/03) (2019) 079.
[45] G. Sparacino, F. Zanderigo, S. Corazza, A. Maran, A. Facchinetti, C. Cobelli, Glucose concentration can be predicted ahead in time from continuous glucose monitoring sensor time-series, IEEE Trans. Biomed. Eng. 54 (5) (2007) 931–937.
[46] C. Zecchin, A. Facchinetti, G. Sparacino, G. De Nicolao, C. Cobelli, Neural network incorporating meal information improves accuracy of short-time prediction of glucose concentration, IEEE Trans. Biomed. Eng. 59 (6) (2012) 1550–1560.
[47] E.H.I. Eliwa, A.M. El Koshiry, T. Abd El-Hafeez, H.M. Farghaly, Utilizing convolutional neural networks to classify monkeypox skin lesions, Sci. Rep. 13 (1) (2023) 14495.
[48] H. Mamdouh Farghaly, M.Y. Shams, T. Abd El-Hafeez, Hepatitis C virus prediction based on machine learning framework: a real-world case study in Egypt, Knowl. Inf. Syst. 65 (6) (2023) 2595–2617.
[49] M.Y. Shams, T. Abd El-Hafeez, E. Hassan, Acoustic data detection in large-scale emergency vehicle sirens and road noise dataset, Expert Syst. Appl. 249 (2024) 123608.
[50] D.A. Abdel Hady, T. Abd El-Hafeez, Predicting female pelvic tilt and lumbar angle using machine learning in case of urinary incontinence and sexual dysfunction, Sci. Rep. 13 (1) (2023) 17940.
[51] D.A. Abdel Hady, T. Abd El-Hafeez, Revolutionizing core muscle analysis in female sexual dysfunction based on machine learning, Sci. Rep. 14 (1) (2024) 4795.
[52] T. Abd El-Hafeez, M.Y. Shams, Y.A. Elshaier, H.M. Farghaly, A.E. Hassanien, Harnessing machine learning to find synergistic combinations for fda-approved cancer drugs, Sci. Rep. 14 (1) (2024) 2428.
[53] A. Omar, T. Abd El-Hafeez, Optimizing epileptic seizure recognition performance with feature scaling and dropout layers, Neural Comput. Appl. 36 (6) (2024) 2835–2852.
[54] D. Javale, S. Desai, Dataset for people for their blood glucose level with their superficial body feature readings, IEEE Dataport (2021).
[55] S.K. Vashist, Continuous glucose monitoring systems: a review, Diagnostics 3 (4) (2013) 385–412.
[56] Y. Guo, G. Bai, Y. Hu, Using Bayes network for prediction of type-2 diabetes, in: 2012 International Conference for Internet Technology and Secured Transactions, IEEE, 2012, pp. 471–472.
[57] D. Pei, T. Yang, C. Zhang, Estimation of diabetes in a high-risk adult Chinese population using j48 decision tree model, Diabetes Metab. Syndr. Obes. (2020) 4621–4630.