Open Journal of Nano
ISSN: 2147-0081
(2024) 9–1
Research Article / Araştırma Makalesi

Blood Glucose Level Estimation Using Photoplethysmography (PPG) Signals with Explainable Artificial Intelligence Techniques

[*1]Gökhan ADIGÜZEL 0009-0004-1545-5427,
[2] Ümit ŞENTÜRK 0000-0001-9610-9550,
[3] Kemal POLAT 0000-0002-7201-6963.
[1] Bolu Abant İzzet Baysal University, Faculty of Engineering, Computer Engineering Depart., Bolu, Turkey.
* Corresponding author, e-mail: gokhanadiguzel1@gmail.com,
[2] Bolu Abant İzzet Baysal University, Faculty of Engineering, Computer Engineering Depart., Bolu, Turkey.
e-mail: umit.senturk@ibu.edu.tr,
[3] Bolu Abant İzzet Baysal University, Faculty of Engineering, Electric and Electronic Engineering Depart., Bolu, Turkey. e-mail: kpolat@ibu.edu.tr.

Submission Date: 24.04.2024 Acceptation Date: 28.04.2024

Abstract - Estimating blood sugar levels is a critical task in effective diabetes management. This study focuses on leveraging the power of machine learning models such as CatBoost, XGBoost, and Extra Trees Regressor, along with explainable AI techniques like SHAP values and confusion matrices, to predict blood sugar levels using Photoplethysmography (PPG) signals. The dataset used in this research is carefully selected for glucose prediction from PPG signals and consists of data from 217 individuals. Information for each individual includes laboratory glucose measurements and approximately one minute of recorded finger PPG signals. Among the various machine learning models tested, CatBoost emerged as the best-performing model in predicting blood sugar levels. The CatBoost model demonstrated its efficiency and accuracy in glucose level predictions by achieving an impressive coefficient of determination (R^2) of 0.71 and a mean absolute error (MAE) of 25.21. Feature importance analysis highlighted the significance of specific features like median deviation and kurtosis in the predictive model built with CatBoost, emphasizing their critical role in determining blood sugar levels. The inclusion of explainable AI techniques enhanced the interpretability and transparency of predictive models. In conclusion, this research underscores the potential of machine learning-based approaches in predicting blood sugar levels from PPG signals. By leveraging advanced models like CatBoost and utilizing explainable AI methods, this study paves the way for improved diabetes management through accurate, non-invasive, and data-driven predictive methodologies.

Keywords: Blood Sugar Prediction, Photoplethysmography, Machine Learning, SHAP, XAI

Açıklanabilir Yapay Zeka Teknikleri ile Fotopletismografi (PPG) Sinyalleri Kullanarak Kan Glikoz Seviyesi Tahmini

Öz - Kan şekeri seviyelerinin tahmini, diyabetin etkili yönetiminde kritik bir görevdir. Bu çalışma, Fotopletismografi (PPG) sinyallerini kullanarak kan şekeri seviyelerini tahmin etmek için CatBoost, XGBoost ve ekstra ağaç regresör gibi makine öğrenimi modellerinin gücünden, SHAP değerleri ve karışıklık matrisi gibi açıklanabilir yapay zekâ teknikleriyle birlikte yararlanmaya odaklanıyor. Bu araştırmada kullanılan veri seti, PPG sinyallerinden glikoz tahmini için dikkatlice seçilmiştir ve 217 kişiden alınan verilerden oluşmaktadır. Her bireyin bilgileri, laboratuvar glikoz ölçümlerini ve yaklaşık bir dakikalık kaydedilen parmak PPG sinyallerini içerir. Test edilen çeşitli makine öğrenimi modelleri arasında CatBoost, kan şekeri seviyelerini tahmin etmede en iyi performansı gösteren model olarak ortaya çıktı. CatBoost modeli, 0.71'lik etkileyici bir determinasyon katsayısı (R^2) metriğine ve 25.21'lik ortalama mutlak hataya (MAE) ulaşarak glikoz seviyesi tahminlerindeki verimliliğini ve doğruluğunu ortaya koydu. Özellik önemi analizi, CatBoost ile oluşturulan tahmin modelinde medyan fark ve basıklık gibi belirli özelliklerin önemini vurgulayarak bunların kan şekeri seviyelerinin belirlenmesindeki önemli rolünün altını çizdi. Açıklanabilir yapay zekâ tekniklerinin dâhil edilmesi, tahmine dayalı modellerin yorumlanabilirliğini ve şeffaflığını arttırdı. Sonuç olarak bu araştırma, PPG sinyallerinden kan şekeri seviyelerinin tahmin edilmesinde makine öğrenimine dayalı yaklaşımların potansiyelini vurgulamaktadır. CatBoost gibi gelişmiş modellerden yararlanan ve açıklanabilir yapay zekâ yöntemlerini kullanan bu çalışma, doğru, invaziv olmayan ve veriye dayalı tahmine dayalı metodolojiler yoluyla gelişmiş diyabet yönetiminin yolunu açıyor.

Anahtar kelimeler: Kan Şekeri Tahmini, Fotopletismografi, Makine Öğrenmesi, SHAP, Açıklanabilir Yapay Zeka

1. Introduction

Diabetes is a significant health issue that affects both individuals and society on a global scale[1]. According to the World Health Organization (WHO), diabetes is considered a major global

DOI: 10.56171/ojn.1473276

health problem and an escalating epidemic[2]. In 2000, the global prevalence of diabetes across all age groups was 2.8%, approximately affecting 171 million people. By 2030, it is estimated to rise to 4.4%, impacting around 366 million people[3]. Diabetes occurs when the pancreas fails to produce enough insulin, or when the body's cells and tissues are unable to effectively use the insulin that is produced[4]. This leads to abnormal levels of glucose in the blood. In healthy individuals, blood glucose levels typically range between 70 and 99 mg/dL. Those with a glucose concentration ranging from 100 to 125 mg/dL are classified as being prediabetic[5]. Refer to Table 1 for the diabetes diagnosis values.[6]

Table 1. Diabetes diagnosis values

Condition	**Fasting Blood Sugar (mg/dL)**	**Oral Glucose Tolerance Test (OGTT) 2 hours after (mg/dL)**	**Hemoglobin A1c (HbA1c) (%)**
Normal	70-99	< 140	< 5.7
Prediabetes	100-125	140-199	5.7-6.4
Diabetes	≥ 126	≥ 200	≥ 6.5

Fasting Blood Sugar (FBS): Blood sugar level measured after at least 8 hours of fasting.

Oral Glucose Tolerance Test (OGTT): Blood sugar level measured 2 hours after consuming 75 grams of glucose.

Hemoglobin A1c (HbA1c): A blood test that indicates the average blood sugar level over the past 2-3 months.

Diabetes can have lasting effects on a person's quality of life. It is a physiological dysfunction characterized by high blood sugar levels due to insufficient insulin, insulin resistance, or excessive glucagon production [7]. Long-term diabetes can lead to chronic complications such as heart disease, kidney disease, stroke, vision loss, and nervous system damage [8]. It can also be a cause of early death for patients with long-term complications. Higher blood sugar levels, known as hyperglycemia, can cause the thickening of blood vessels, leading to kidney damage, vision loss, and sometimes even organ failure. Diabetes is also associated with limb amputation, peripheral vascular diseases, and myocardial issues.

Today, both invasive and non-invasive methods are used to measure blood sugar. These methods describe different approaches for medical diagnosis or monitoring. Invasive Measurement Method: Invasive measurement methods involve gathering information by entering the body or penetrating the skin. These methods typically offer direct access to the patient's internal organs and are used in various medical procedures. For instance, cardiac catheterization is commonly used for diagnosing heart diseases and conducting surgical procedures, while blood sampling is a standard practice for laboratory tests. In general, invasive measurement methods provide more accurate results by directly accessing the target organ or tissue inside the body. As a result, they are preferred for the diagnosis and treatment of critical illnesses. Non-Invasive Measurement Method: Non-invasive measurements gather information without penetrating the body or with minimal contact. Patient preferences often lean towards less invasive non-invasive methods due to the potential risks and discomfort associated with invasive methods, especially for the monitoring of chronic diseases. Non-invasive methods are generally preferred for monitoring chronic diseases because they are more frequently used and less invasive. Non-invasive measurement methods generally provide less accurate results compared to invasive methods because they collect information from outside the body. However, advancements in technology are improving the accuracy of non-invasive methods.

Both invasive and non-invasive measurement methods require physical contact. However, non-invasive methods are limited to the skin surface, causing less discomfort than invasive methods. Patients often prefer non-invasive methods due to their lower invasiveness, reduced risk, and increased comfort. Factors such as patient comfort, lower costs, and reproducibility contribute to the popularity of non-invasive methods. However, the severity of the disease and specific monitoring parameters may sometimes necessitate the use of invasive methods. Table 2 displays the features of invasive and non-invasive blood glucose measurement methods currently in use.

DOI: 10.56171/ojn.1473276

Open Journal of Nano
ISSN: 2147-0081
(2024) 9–1
Research Article / Araştırma Makalesi

Table 2. Invasive and non-invasive blood glucose measurement methods

Method	Description	Invasive/Non-invasive
Fingerstick Testing (Blood Glucose Test) [6]	Uses a needle or lancet to collect a blood sample from the fingertip, which is measured with a glucose meter.	Invasive
Continuous Glucose Monitoring (CGM) [9]	A sensor placed under the skin continuously measures blood glucose levels and transmits data to a monitor or smartphone.	Non-invasive
Intravenous Blood Sampling [10]	Blood samples are taken intravenously for laboratory tests and analysis.	Invasive
Electrochemical Methods [11]	Uses electrodes to measure glucose levels in a blood sample. Glucose reacts with an enzyme, producing an electrical signal proportional to glucose amount.	Invasive
Transdermal Glucose Monitoring [12]	A device placed on the skin measures glucose levels through the skin using technologies like electromagnetic waves or ultrasound.	Non-invasive
Saliva or Tear Analysis [13]	Measures glucose levels in body fluids like saliva or tears, still in research and development stages.	Non-invasive.
Optical Sensors[14]	Uses optical sensors, like laser light or infrared light, to determine glucose levels.	Non-invasive
Bioelectrical Impedance Analysis[15]	Measures electrical properties of the body to determine glucose levels with electrodes placed on the skin's surface.	Non-invasive

Non-invasive methods are preferable for continuous glucose monitoring compared to invasive and semi-invasive methods[16]. Optical methods have been found to be more reliable and accurate in glucose measurement[17]. Common optical methods for non-invasive measurements include Raman spectroscopy, near-infrared spectroscopy, polarimetry, scattering spectroscopy[18], photoacoustic spectroscopy [19], and others. Non-invasive measurements can address previous issues and provide painless and accurate solutions [16]. By using a non-invasive method, blood sampling is eliminated, reducing the risk of bacterial and viral infections, as well as alleviating patient discomfort. The following are the application steps for invasive and non-invasive methods.

Traditional Blood Glucose Measurement System: This system is invasive and is generally used as part of the initial insulin-based treatment. These types of devices are typically glucose monitors used in conjunction with active medication therapy and rely on advanced optical sensor technology. Optical sensor technology, applied solely for the detection of glucose, allows for the direct acquisition of data. The primary purpose of a glucose monitor is to detect the maximum glucose level in the body. However, since these monitors can increase the risk of infections, there is a need for alternative methods. For patients to measure their blood sugar levels at home, they require the devices and equipment shown in Figure 1. Once the devices and equipment are obtained, the patient begins a difficult and sometimes painful process. The traditional blood sugar measurement procedure is usually not something the patient can do alone.

Figure 1: Blood glucose measurement device and equipment.

Precautions need to be taken both before and after the measurement. The finger from which the measurement will be taken must be kept sterile to prevent infection. The stages of blood sugar measurement at home are shown in Figure 2 below.

DOI: 10.56171/ojn.1473276

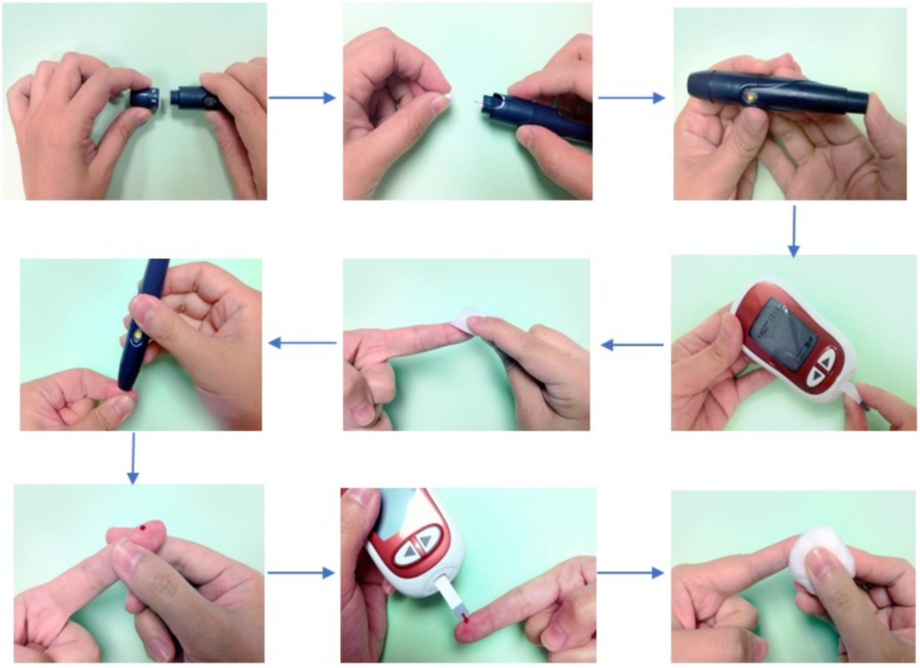


Figure 2. Stages of blood sugar measurement at home.

Traditional methods for determining blood glucose levels include using subcutaneous measurements and invasive interventions. In cases where continuous monitoring is necessary, a sensor is implanted under the patient's skin to constantly track glucose levels. These sensors continuously take and record measurements. This procedure is typically carried out in a hospital setting under the supervision of a physician. Research on electrochemical measurement systems aims to reduce sensor sizes and improve measurement accuracy. In a study conducted by Kubilay A., Hakan B., and their colleagues, a reusable and enzyme-free sensor was developed. The absence of enzymes in these sensors helps to reduce potential complications in the body. [20]

In a study by Fatemeh Karimi and colleagues, economical production methods were used to create enzyme-free sensors[21]. The developed sensor was found to be extremely stable, discriminative, and sensitive, indicating its potential as a glucose sensor for the market. These sensors are designed to overcome the disadvantages of traditional glucose sensors and invasive methods, offering the potential to improve glucose detection systems in the future.

In addition to electrochemical measurements, it's important for patients to measure blood sugar without invasive interventions. Continuous measurements are necessary to prevent the disease from reaching severe levels. Although the size of interventional method sensors has been reduced and enzyme-free sensors have been developed, traditional methods are uncomfortable for the patient and time-consuming. Therefore, the development of noninvasive blood glucose measurement systems is crucial. Measures can be taken to achieve noninvasive blood sugar measurement in the least painful and shortest time by utilizing signals received from the body. Photoplethysmography (PPG) sensors are used to measure these signals. These sensors work by using light to measure blood flow and volume and are often used to monitor heart rate, blood oxygen saturation (SpO2) and other cardiovascular parameters [22]. The signal measurement system of the PPG sensor is shown in Figure 3. The potential of artificial intelligence is revealed when analyzing PPG signals. As shown in Figure 4, signals measured with the PPG sensor require interpretation that involves a complex set of mathematical operations. This process is difficult and cumbersome to accomplish with human intelligence and effort alone.

Open Journal of Nano
ISSN: 2147-0081
(2024) 9–1
Research Article / Araştırma Makalesi

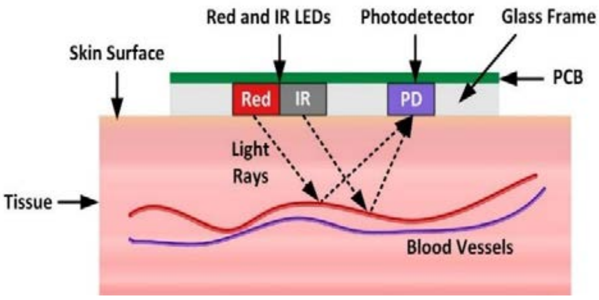


Figure 3: PPG signal measurement system

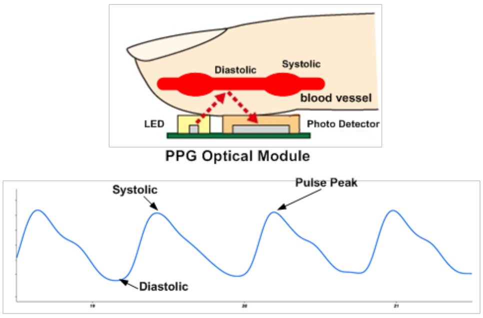


Figure 4: PPG sensor signal

With the advancement of AI algorithms, it is now feasible to improve the accuracy of non-invasive blood sugar measurements. This study seeks to gauge blood sugar levels without invasive procedures, such as drawing blood. The use of non-invasive systems is expected to provide unbiased data, potentially enhancing the effectiveness of diabetes treatment.

2. Materials and Methods

The methodological steps in the study were started with the data set being read and the data being preprocessed. The signals were first passed through a filter. This step usually involves the data being cleaned, noise being removed, and the signals being converted into a format suitable for further processing. After filtering, the cleaned signals were subjected to feature extraction. This step involves relevant features or characteristics being identified and extracted from the signals that are informative and useful. In the next phase, the training phase, the relationship between the features and the target variable was learned by the model. Finally, the final output of the process was a trained regression model. This model was used to predict blood glucose levels based on new input signals using the relationships learned during the training phase. These steps are illustrated in Figure 5.

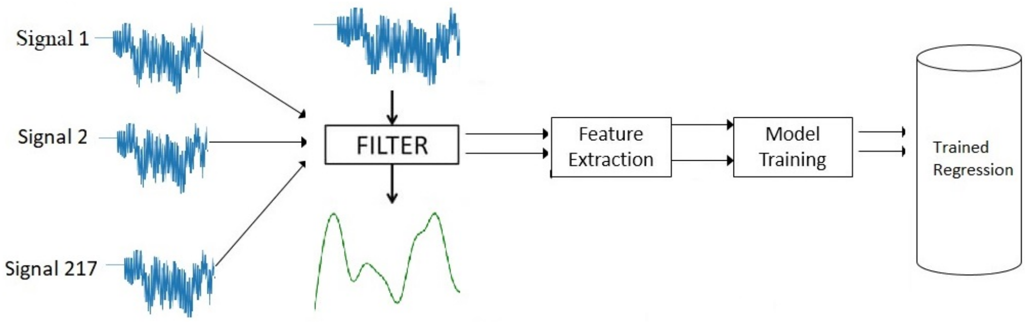


Figure 5: Workflow Steps

2.1. Dataset

In the development of the proposed system, the database in the literature was taken into account [23]. This database contains glucose values from 217 participants (127 women and 90 men) and PPG signal values from a clinical-chemical and immunology analysis unit (Cobas 6000) at a hospital in Cuenca, Ecuador. The PPG signal was collected for an average of 2.5 minutes at a sampling frequency of 64 Hz using the Empatica E4 wristband. The wristband uses green (550 nm) and red (650 nm) lights to optimize pulse wave detection [24]. The glucose value obtained in the laboratory environment was also assigned as the target variable.

2.2. Data Preprocessing

The first 10 seconds of data were removed from the signal to avoid errors due to device initial parameters. Signal preprocessing was performed using the NeuroKit2 library [25] and the Elgendi [26] method was used to remove noise and artifacts in PPG signals. In Figure 6, the Elgendi method,

DOI: 10.56171/ojn.1473276

Open Journal of Nano
ISSN: 2147-0081
(2024) 9–1
Research Article / Araştırma Makalesi

known as broadband amplitude integration, was used to display the raw signal and the cleaned signal. The Elgendi method is used to evaluate the broadband characteristics of the PPG signal. Additionally, this method is effective in improving the quality of the PPG signal and integrating amplitudes into specific frequency bands.

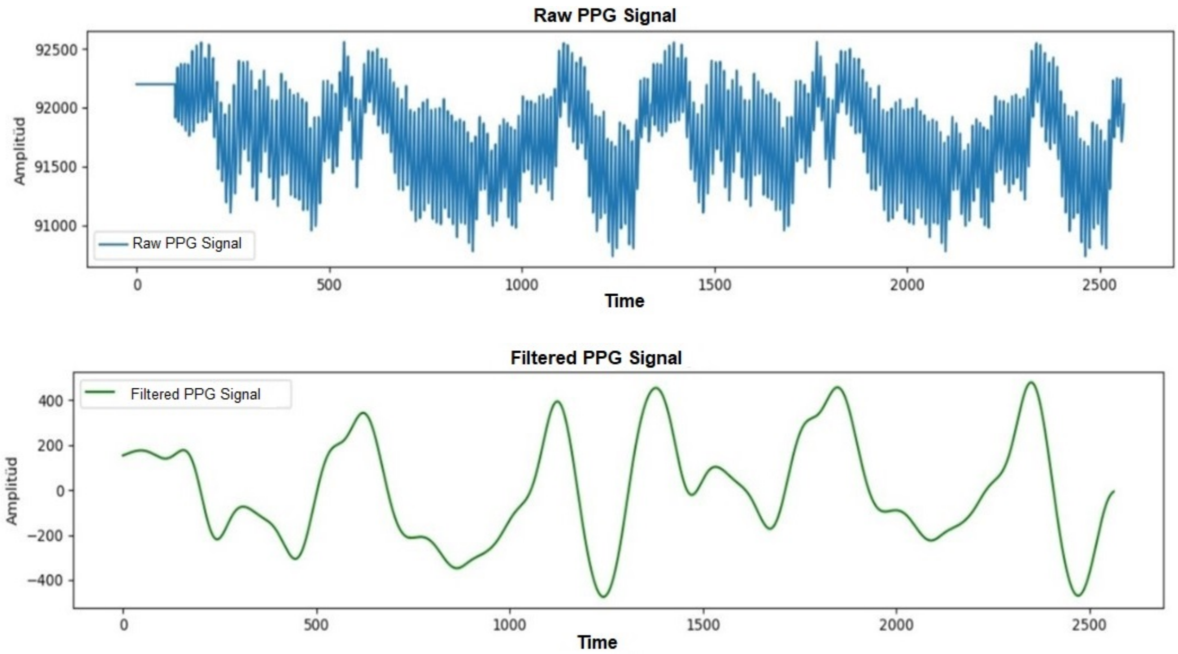


Figure 6: Signal Processing Steps

The PPG signal is obtained by detecting light that is reflected or absorbed from tissues and blood vessels in the body. This signal is commonly used in pulse measurements and medical monitoring applications. The Elgendi method is employed in the analysis of pulse waves in PPG signals, aiming to determine the frequency, amplitude, and phase changes of the pulse waves. This method relies on a customized wavelet transform technique to extract prominent features of the pulse waves. The Elgendi method is a widely used and effective technique in processing PPG signals. It plays a significant role in various application domains such as medical diagnosis, health monitoring, and sports performance monitoring, as it allows for the extraction and analysis of prominent features of pulse waves.

2.3. Segmentation

The segmentation of a Photoplethysmography (PPG) signal involves dividing it into specific events or periods. This signal typically represents the heart cycle, and segmentation is used to identify particular events or features within the cycle. Various methods can be used for PPG signal segmentation, such as R-peak detection, inverted peak detection, start and end of heartbeats, and frequency domain analysis. In this study, PPG signals were segmented based on the start and end intervals of heartbeats, as shown in Figure 7.

DOI: 10.56171/ojn.1473276

Open Journal of Nano
ISSN: 2147-0081
(2024) 9–1
Research Article / Araştırma Makalesi

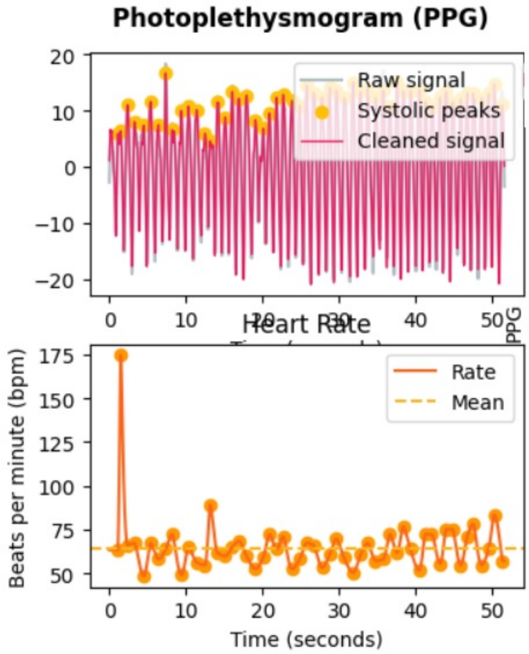


Figure 7: Segmentation Process

Segmented PPG signals were averaged to make it easier to analyze specific events or conditions. Averaging is particularly useful when working with repeated cycles, such as a PPG signal, as it allows us to identify the characteristics of each cycle. Figure 8 showed the starting and ending values as well as the average value of the heartbeat of patient number 3. The dark red line represents the average heart rate, and the light thin lines represent all heartbeats of the patient.

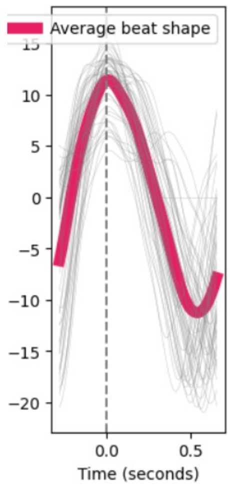


Figure 8: Segmentation Processy

2.4. Feature Extraction and Selection

PPG signals were preprocessed using the TSFEL[27] library, a Python library designed to extract various features from time series data, and then converted them into features. The library provides a wide range of feature extraction functions and capabilities for generating a feature matrix. Available features are divided into three categories: statistical, temporal and spectral. In the study, temporal and statistical features were used to divide the PPG signal into 230 features. Some features, such as the maximum value of a signal, ECDF value calculated over time (empirical cumulative distribution function), and signal entropy, were excluded from the study because they did not

Open Journal of Nano
ISSN: 2147-0081
(2024) 9–1
Research Article / Araştırma Makalesi

contribute to blood glucose prediction. A total of 51 signal features were ultimately used for the prediction calculation, and Table 3 contains a list of some of the features used.

Table 3. Some of the Features to be Used

PPG Features	Description
1	Peak value
2	Peak-to-peak
3	Mean value
4	Standard deviation
5	Root Mean Square (RMS)
6	Skewness
7	Kurtosis
8	Crest factor
9	Impulse factor

1- Peak Value: It represents the highest value of a time series. This value determines the peak of the time series and represents the maximum value that occurs over a certain period. Identifying the highest value in time series is especially important in areas such as anomaly detection, signal processing, and data mining.[28]
2- Peak-to-Peak: It refers to the difference between the highest (peak) and lowest (trough) points of a time series. This metric is used to measure the fluctuation amplitude of the signal and represents the variability or fluctuation range of the time series.[28]
3- Mean Value: It is the value obtained by dividing the sum of the elements of a time series or data set by the number of elements. The mean is used as a measure of central tendency and represents the overall trend or typical value of the data set.[28]
4- Standard Deviation: It is a statistical value that measures how much the elements in the data set deviate from the average. Standard deviation is used to understand the spread or dispersion of a data set and shows how the data is distributed around the mean.[28]
5- Root Mean Square(RMS): It is a measure defined as the square root of the mean square value of a data set. RMS is used primarily in fields such as signal processing, engineering, and statistics and represents the energy content or amplitude level of a data set.[28]
6- Skewness: It is a statistical measure that describes the asymmetry of data distribution around the mean. Indicates whether data points are denser on one side of the average than the other. Skewness can be positive, negative, or zero, and each case tells us something different about the shape of the data distribution.[29]
7- Kurtosis: It is a statistical measure that describes the shape of a distribution's tails about its overall shape. It indicates whether the data points in the tails of the distribution are heavy or light compared to a normal distribution. In other words, kurtosis measures the "tailedness" or the extremity of deviations in a dataset.[29]
8- Crest Factor: It is a measure expressing the ratio of the peak value of a signal or data set to the average value. It is often used in fields such as electrical engineering, audio engineering, and signal processing. The crest factor determines the relationship between the peak value and average value of a signal and provides information about the amplitude of the signal.[30]
9- Impulse Factor: This is the ratio of the highest amplitude (peak) value of a signal to the duration of the signal. Impulse factor is used to describe the pulse-shaped characteristic of a signal and provides information about the overall performance of the signal.[28]

The signals, separated according to the determined features, were associated with the glucose values obtained in the laboratory as the target variable and prepared for the machine learning model.

2.5. Machine Learning Model

PPG signals were subjected to preprocessing and feature extraction before being regressed with the target variable. Regression involves modeling and understanding the relationship between

DOI: 10.56171/ojn.1473276

variables, which is commonly used to analyze the relationship between the dependent variable (outcome variable) and the independent variables (predictor variables). In our research, regression models in the Pycaret [31] library were used to model the relationship between the target variable and one or more independent variables. These models help us understand how a variable depends on other variables and predict future values. We used a series of regression models to increase training depth. The performance values of the first seven regression models are presented in Table 4, and the performance values of all regression models are detailed in the results section.

Table 4. Performance Values of the First Seven Regression Models

Model	MAE	MSE	RMSE	R^2	RMSLE	MAPE	TT(Sec)
CatBoost Regressor	25.162	1591.7898	39.1387	0.7191	0.2436	0.1837	31.150
Light Gradient Boosting Machine	28.6411	1941.9567	43.4559	0.6553	0.2692	0.2029	0.0250
Extra Trees	26.9872	2095.2192	44.5275	0.6428	0.2790	0.1996	0.0770
Gradient Boosting Regressor	29.9443	2368.1302	47.9978	0.5960	0.2870	0.2084	0.0880
AdaBoost Regressor	36.7854	2400.9081	48.2477	0.5890	0.2951	0.2612	0.0420
Random Forest Regressor	29.5800	2493.4937	48.9901	0.5693	0.2932	0.2089	0.1030
Extreme Gradient Boosting	30.9313	2833.0924	52.1906	0.5263	0.3067	0.2164	0.2100

The table includes various metrics that measure the performance of different machine-learning models. For each model, mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE), coefficient of determination (R^2), root mean square logarithmic error (RMSLE), mean absolute percentage error (MAPE) and training It is seen that it is evaluated with metrics such as duration (TT). MAE (Mean Absolute Error): Indicates the average absolute difference between the actual values and the predicted values by the model. A lower value indicates better prediction performance [32]. MSE (Mean Squared Error): Indicates the average squared difference between the actual values and the predicted values by the model. A lower value indicates better prediction performance [33]. RMSE (Root Mean Squared Error): Obtained by taking the square root of MSE. It represents the square root of the average squared difference between the actual and predicted values. A lower value indicates better prediction performance [34]. R^2 (R-squared): Indicates the percentage of variation in the dependent variable explained by the independent variables. Values range from 0 to 1, where a higher value indicates a better fit of the model [35]. RMSLE (Root Mean Squared Logarithmic Error): Represents the square root of the average squared difference between the logarithms of the actual and predicted values. Particularly useful when there are large values in the dataset [36]. MAPE (Mean Absolute Percentage Error): Represents the average absolute percentage difference between the actual and predicted values. A lower value indicates better prediction performance [37]. TT (Training Time): Represents the training time of the model. A shorter time indicates faster training time. In the trained system used in the study, two models with the highest performance among regression models were analyzed: CatBoostRegressor [38] and LightGBM (Light Gradient Boosting Machine) [39]. CatBoost directly supports categorical variables in regression tasks. This allows users to use categorical variables without the need for special handling or conversion. CatBoost is based on the gradient boosting algorithm. Gradient boosting aims to create a strong model by combining weak learners, usually decision trees. LightGBM (Light Gradient Boosting Machine) provides fast training on large datasets and datasets with high-dimensional features by using parallel processing and a specialized histogram-based learning method.

2.6. Explainable Artificial Intelligence (SHAP)

In this study, Feature Importance Graph was used instead of performance data. Figure 9 shows the Feature Importance Graph of the CatBoostRegressor algorithm, which achieved the highest performance metric. A Feature Importance Graph is a graph that visualizes the contribution of features (variables or attributes) in a machine learning model to the model's performance. It helps understand the impact on the model's output by ranking each feature in order of importance.

Open Journal of Nano
ISSN: 2147-0081
(2024) 9–1
Research Article / Araştırma Makalesi

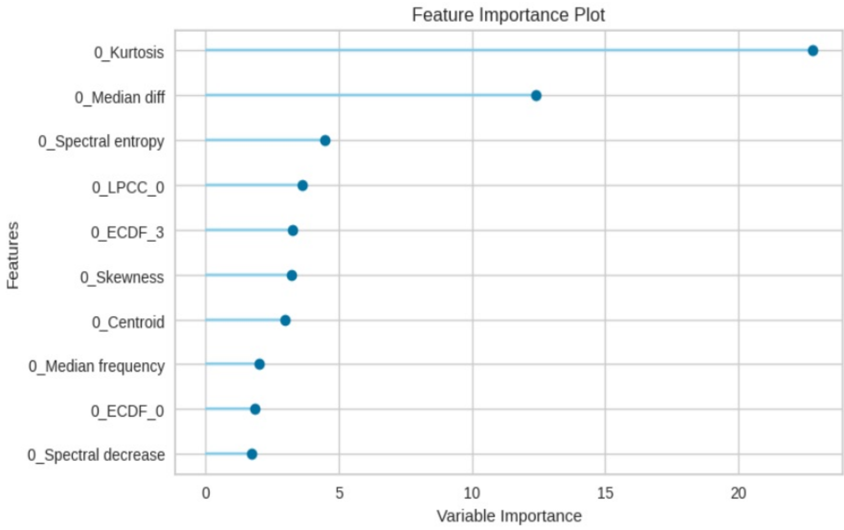
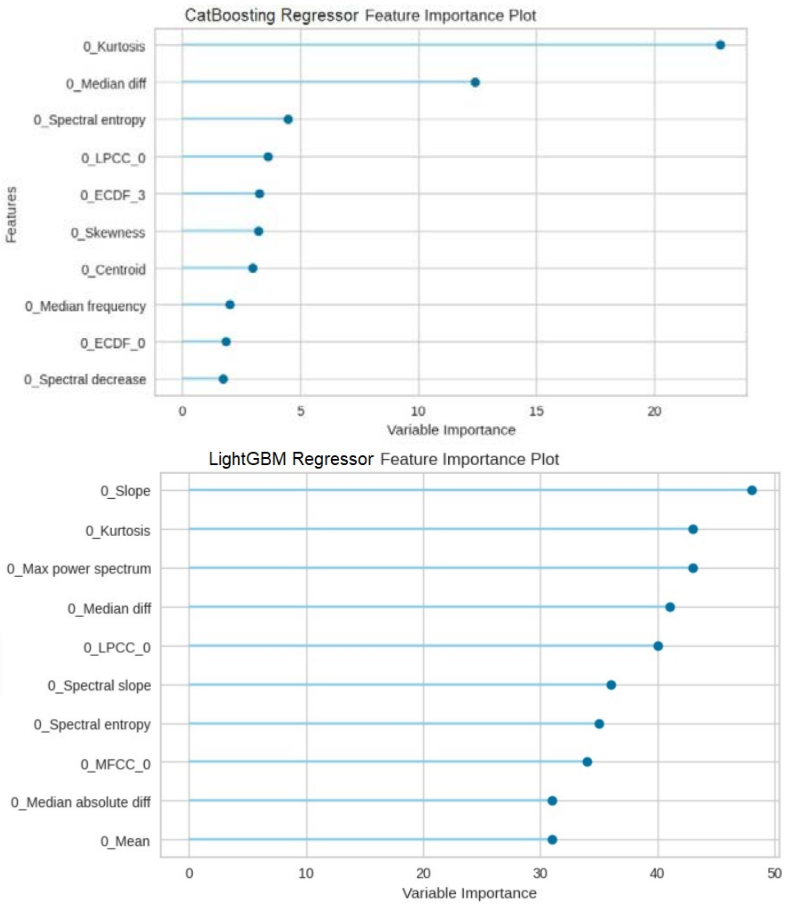


Figure 9: Feature Importance Plot for the CatBoostRegressor Algorithm

In the study, the performance of the Light Gradient Boosting Machine algorithm, which is the second high-performance algorithm used, was analyzed. In examinations of both training and testing data, the algorithms showed similar performance metrics. However, differences were observed in the Feature Importance Graph. Comparison of the CatBoostRegressor algorithm with the Light Gradient Boosting Machine algorithm revealed differences in feature ranking between the two regression algorithms. These differences in feature importance rankings are shown in Figure 10.



Open Journal of Nano
ISSN: 2147-0081
(2024) 9-1
Research Article / Araştırma Makalesi

Figure 10: CatBoostRegressor - Light Gradient Boosting Machine Feature Importance Plot Comparison

The two graphs show the varying cardinality of the CatBoost and LightGBM regression algorithms. Both algorithms work on a specific data set to determine which features are more important when predicting the target variable.

Kurtosis: It is seen as a very important feature in both models. Kurtosis formulas given formula 1. This suggests that this feature plays a critical role when predicting the target variable in the dataset.
Median Diff: While it is very important in the CatBoost model, it ranks fourth in the LightGBM model. This difference may be due to the different algorithmic structures of the two models. Slope: While it is considered the most important feature in LightGBM, it is not even among the top 10 features in the CatBoost model. This may give a clue as to how features are evaluated in different models.
Other Features: Different features stand out in both models. For example, while 0_Spectral entropy was found to be significant in both models, 0_Spectral decrease was only significant in the CatBoost model. This analysis is useful for understanding which features are consistently important across both models and which features vary from model to model. This information can be taken into account during model selection and improvement.

The Kurtosis [40] value, the most important determining feature in the CatBoostRegressor algorithm, is a measure that represents how peaked or flat a data set is, typically compared to a normal distribution. It indicates the degree to which data points are more or less sharply distributed.

$$\text{Kurtosis} = \frac{\mathrm{n} * \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{i}\Sigma\mathrm{ni}(\mathrm{Yi} - \bar{\mathrm{Y}})^4}{\left(\sum_{i=1}^{n} i(\mathrm{Yi} - \bar{\mathrm{Y}})^2\right)^2} \tag{1}$$

Yi: Distribution Variable i
Ȳ: Mean of the Distribution
n: Number of Variables in the Distribution

In the Light Gradient Boosting Machine algorithm, the most significant determining feature is the Slope [40] function. The Slope function calculates the slope of a data series. The slope represents the rate of change in a data series and indicates the steepness of a line or curve in two-dimensional space.

$$y = ax + b,$$

y = Dependent Variable
a = Coefficient
x = Independent Variable
b = Constant Value

2.7. Blood Glucose Level Prediction

The study aimed to evaluate the performance of various regression models and select the best performers for further analysis. After evaluating different metrics, the regression models with the highest performance were determined. These models were then retrained using the dataset to ensure optimal performance. A “prediction error” plot was used to gain insight into the performance of the selected models and to understand how well they predicted the target variable. This type of chart is commonly used in data analysis and machine learning to visualize disparities between actual values

DOI: 10.56171/ojn.1473276

Open Journal of Nano
ISSN: 2147-0081
(2024) 9–1
Research Article / Araştırma Makalesi

and predicted values generated by the model. In general, the prediction error chart serves as a visual tool to evaluate the performance of regression models and identify areas for improvement. It helps researchers and analysts make informed decisions about the reliability and effectiveness of models in predicting the target variable.

3. Results and Discussions

The study compared regression models trained with relevant settings and parameters in the data. The performance metrics used were Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Square Error (RMSE), R-squared (R^2), Root Mean Square Logarithmic Error (RMSLE), and Mean Absolute Percentage Error. (MAPE). Performance metrics for all regression models are shown in Table 5.

Table 5. Performance Metrics Based on Regression Models

Model	MAE	MSE	RMSE	R^2	RMSLE	MAPE	TT(Sec)
CatBoost Regressor	25.16	1591.78	39.13	0.71	0.24	0.18	31.15
Light Gradient Boosting Machine	28.64	1941.95	43.45	0.65	0.26	0.20	0.02
Extra Trees	26.98	2095.21	44.52	0.64	0.27	0.19	0.07
Gradient Boosting Regressor	29.94	2368.13	47.99	0.59	0.28	0.20	0.08
AdaBoost Regressor	36.78	2400.90	48.24	0.58	0.29	0.26	0.04
Randorn Forest Regressor	29.58	2493.49	48.99	0.56	0.29	0.20	0.10
Extreme Gradient Ba)sting	30.93	2833.09	52.19	0.52	0.30	0.21	0.21
Lasso Regression	38.30	2967.40	537.92	0.48	0.30	0.24	0.01
Ridge Regression	37.09	2795.83	521.44	0.47	0.30	0.24	0.08
Bayesian Ridge	39.49	3075.62	546.58	0.46	0.31	0.25	0.01
Lasso Least Angle Regression	39.93	3209.94	561.38	0.43	0.32	0.26	0.01
Decision Tree Regressor	31.58	3571.12	572.14	0.42	0.33	0.20	0.01
Elastic Net	43.97	3660.27	597.49	0.34	0.34	0.29	0.01
Linear Regression	41.37	3537.18	578.18	0.29	0.37	0.28	0.01
K Neighbors Regressar	43.58	4315.71	649.02	0.20	0.37	0.31	0.01
Orthogonal Matching Pursuit	60.01	5908.77	757.43	0.01	0.44	0.41	0.01
Huber Regressor	60.40	7068.74	820.82	-0.09	0.51	0.36	0.01
Dummy Regressor	70.61	7049.36	834.95	-0.18	0.48	0.50	0.0070

A lower MSE indicates better prediction performance. The best-performing model in terms of MSE is the "CatBoost Regressor" model(1591.78). A lower RMSE indicates better forecast performance. Among the models in the table, the best-performing model in terms of RMSE is the "CatBoost Regressor" model(39.13). If the R^2 value, which shows how much of the variance of the dependent variable is explained by the independent variables, is close to 1, it means a better fit. Among the models in the table, the best-performing model in terms of R^2 is the "CatBoost Regressor" model (0.71). A lower RMSLE, which represents the square root of the logarithmic error amount, indicates a better prediction performance. Among the models in the table, the best-performing model in terms of RMSLE is the "CatBoost Regressor" model (0.24). A lower MAPE value, which represents the average absolute percentage difference between actual values and predicted values, indicates a better prediction performance. Among the models in the table, the best-performing model in terms of MAPE is the "CatBoost Regressor" model (0.18). A lower training time in the Training Time value, which represents the time it takes to train the model, indicates faster model training. Among the models in the table, the best-performing model in terms of training time is the "CatBoost Regressor" model (31.150 seconds).

The primary performance metric used in the study is the R-squared metric. R-squared is a statistical measure used to evaluate the performance of regression models. The R-squared value indicates how well the dependent variable is explained by the independent variables. This measure ranges from 0 to

DOI: 10.56171/ojn.1473276

1, with values closer to 1 indicating a better fit, while values approaching 0 indicate a weaker fit. R-squared is often expressed by the formula 2;

$$R^2 = 1 - \frac{\sum i{=}1n(yi-y\hat{}i)2}{\sum i{=}1n(yi-y\hat{}i)2} \qquad (2)$$

yi = Actual Values, $y\hat{}i$ = Predicted Values, $y\bar{}$ = Mean of the Dependent Variable, $\sum i{=}1n$ = Sum , Number of Observations.

In the prediction error plot for the R2 metric, the x-axis represents the actual values and the y-axis represents the predicted values. Each point represents the true value of an observation and the model's prediction for that observation. Ideally the points should be aligned along the 45-degree line. A skewed distribution or irregular patterns in such charts may indicate that the model is making an error on certain data points. The R2 prediction error graph of the CatBoostRegressor model, which has the highest performance, is shown in Figure 11. The R2 performance metric value was found to be 0.71.

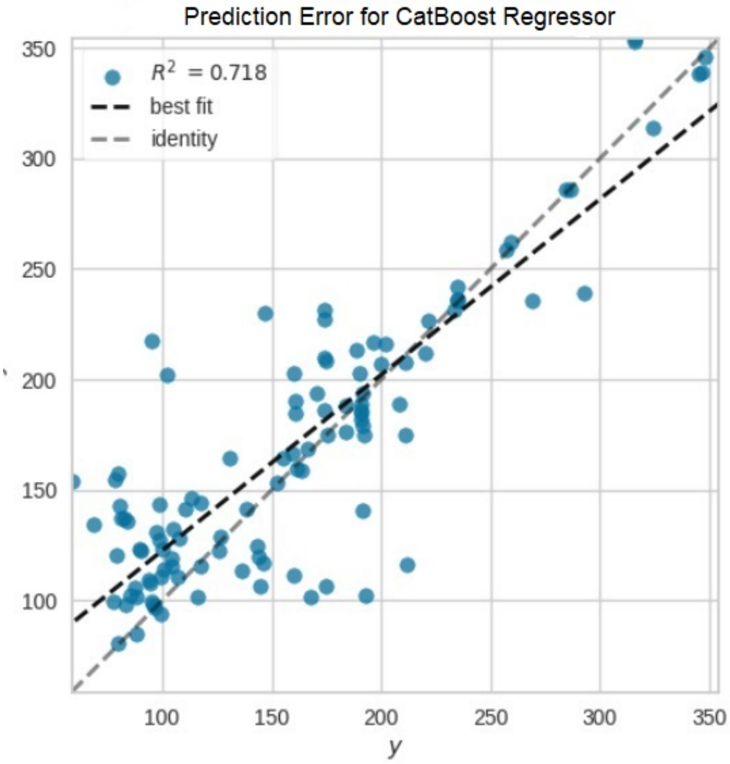


Figure 11: Performance Metrics Based on Regression Models

Below, Table 6 displays five studies from the literature.

Table 6. Five studies from the literature.

Reference	Subject of Study	Classifier	Performance
Geshwaree Huzooree et al. – 2017[41]	Glucose Prediction Data Analytics for Diabetic Patients Monitoring	Autoregressive (ARX)	R^2 = 0.342
Guillermo Edinson et all. – 2020 [42]	Application of Artificial Intelligence Techniques for the Estimation of Basal Insulin in Patients with Type I Diabetes	NNs and BNs	NNs R^2 = 0 . 067 BNs R^2 = 0. 634
Nanayakkara et al. - 2018 [43]	Non-Invasive Blood Glucose Monitoring using a Hybrid Technique	BIS + NIRS	R^2 = 0.58

Open Journal of Nano
ISSN: 2147-0081
(2024) 9–1
Research Article / Araştırma Makalesi

Rui Sun et al. – 2023 [44]	Time in Range Estimation in Patients with Type 2 Diabetes is Improved by Incorporating Fasting and Postprandial Glucose Levels	Linear Regression	R^2 = 0.36
Chowdhury Azimul Haque et al. – 2020 [45]	Noninvasive In Vivo Estimation of Blood-Glucose. Concentration by Monte Carlo Simulation	XGBoost	R^2 = 0.68
Pappada, S. M. et al. - 2011 [46]	Neural network-based real-time prediction of glucose in patients with insulin-dependent diabetes.	Neural Network	RMSE= 43.9
Aishah, A. F. Q. A. et al. -2011 [47]	Multiple linear regression model analysis in predicting fasting blood glucose level in healthy subjects.	Multiple linear regression	R^2=0.59
Our proposed method		**CatBoost Regressor**	**R^2=0.71,** **RMSE= 39.13**

Table 6 presents a comparative analysis of several studies on glucose prediction and monitoring in diabetic patients, along with a proposed model. Geshwaree Huzooree et al. (2017) used an autoregressive (ARX) model, achieving an R^2 of 0.342, indicating moderate accuracy. Guillermo Edinson et al. (2020) applied AI techniques, with Neural Networks (NNs) performing poorly (R^2 = 0.067) and Bayesian Networks (BNs) performing well (R^2 = 0.634). Nanayakkara et al. (2018) utilized a hybrid technique (BIS + NIRS) with an R^2 of 0.58, showing good accuracy. Rui Sun et al. (2023) employed linear regression for time-in-range estimation, yielding an R^2 of 0.36. Chowdhury Azimul Haque et al. (2020) used XGBoost, achieving the highest R^2 of 0.68 among these studies. Pappada et al. (2011) used a neural network with an RMSE of 43.9, and Aishah et al. (2011) used multiple linear regression with an R^2 of 0.59. The proposed model, using a CatBoost Regressor, outperforms all others with an R^2 of 0.71 and an RMSE of 39.13, indicating superior predictive accuracy and lower error.

4. Conclusions

Machine learning is an effective method for predicting blood sugar levels based on PPG signals. PPG is a type of signal obtained from the skin through the absorption or reflection of light, which changes according to the heartbeat. These signals reflect changes in blood circulation and provide information about glucose levels in the blood. Machine learning algorithms can analyze the complexity of PPG signals and predict individuals' blood sugar levels using data from these signals. This method can be considered an alternative to traditional blood sugar measurements because it works noninvasively and can improve individuals' quality of life. However, more research and development are needed to ensure the reliability and accuracy of such systems.

When the success of this study was compared with similar studies, it was seen that the success rate was high. It is possible to increase the success rate by making adjustments such as parameter setting and reducing the noise characteristic during signal measurement. This study used PPG signals to estimate blood sugar levels. Signals were preprocessed to remove noise and artifacts, followed by segmentation based on peaks. The segmented signals were then averaged to prepare for feature extraction. Feature extraction aimed to identify and represent important information from the dataset, incorporating statistical and temporal methods. Specifically, 160 statistical features were extracted and 51 features were selected based on their relevance to the target variable. The pre-processed, segmented, and feature-separated signals were then trained using machine learning models. The system demonstrated success with regression models, achieving an R-squared performance measure of 0.71.

DOI: 10.56171/ojn.1473276

This study focused on explainability and interpretability using explainable artificial intelligence techniques (SHAP) to analyze high-performance models. Feature Importance Plot was used to visualize the contribution of features to the performance of the model. It has been noted that the ranking of features that have the most significant impact on performance varies between machine learning models. The study concluded that analyzing signal properties not only with statistical and temporal attributes, but also with frequency domain attributes can positively affect the results. Additionally, it was stated that increasing the number of samples will significantly affect the feature results. It has been suggested that the small number of extracted features that are highly relevant to the target variable will be important for future Internet of Things (IoT) studies and will enable integration into embedded systems.

Peer-review: Externally peer-reviewed.

Conflict of Interest: This paper has been presented at the INERS'2024 (5th International Engineering Research Symposium) held in Düzce (Turkey), March 7-9, 2024.)

Financial Disclosure: The authors declared that this study has received no financial support.

References

[1] International Diabetes Federation. (IDF 2009c). Diabetes Prevalence. Online Access

http://www.idf.org/home/index.cfm?node=264 (Acces:12.05.2009)

[2] Diyabet Çığ Gibi Büyüyor. (2008). Diyabete Bakış, 7, 6-7.

[3] K. Wikblad, L. Wibell, , K. Montin (1990). “The Patient's Experience Of Diabetes And Its Treatment: Construction Of An Attitude Scale By A Semantic Differential Technique.” Journal Of Advanced Nursing, 15(9), 1083-1091.

[4] American Diabetes Association. Diagnosis and classification of diabetes mellitus, Diabetes Care, 32 (1), 62–67, 2009.

[5] O. Sevli. Diyabet hastalığının farklı sınıflandırıcılar kullanılarak teşhisi, Journal of the Faculty of Engineering and Architecture of Gazi University, 38:2 (2023) 989-1001

[6] American Diabetes Association. (2023). Standards of Medical Care in Diabetes—2023. Diabetes Care, 46(Suppl. 1): S1-S291.

[7] Lam, B. Q., Srivastava, R., Morvant, J., Shankar, S., Srivastava, R. K. 2021. Association of DM and Alcohol Abuse with Cancer: Molecular Mechanisms and Clinical Significance. Cells 2021, Vol. 10, 10(11):, 3077. https://doi.org/10.3390/CELLS10113077

[8] Quinones, S., Robert Roberts, C. C., David Cistola, C.-C., Narayan, M., Crites, S. L. 2021. Non-invasive in-vitro glucose monitoring using an optical sensor and machine searning techniques for diabetes applications, Texas.

[9] Beck, R. W., Riddlesworth, T., Ruedy, K., & Kollman, C. (2018). Continuous glucose monitoring versus usual care in patients with type 2 diabetes receiving multiple daily insulin injections: a randomized trial. Annals of Internal Medicine, 169(6), 379-387.

[10] Karon, B. S. (2016). Why is there still no international standard for glucose in blood? Clinical Chemistry and Laboratory Medicine (CCLM), 54(6), 975-977.

[11] Newman, J. D., & Turner, A. P. (2005). Home blood glucose biosensors: a commercial perspective. Biosensors and Bioelectronics, 20(12), 2435-2453.

Open Journal of Nano
ISSN: 2147-0081
(2024) 9–1
Research Article / Araştırma Makalesi

[12] Sieg, A., Guy, R. H., & Delgado-Charro, M. B. (2012). Non-invasive and minimally invasive methods for measuring glucose. In Advances in Noninvasive Electrocardiographic Monitoring Techniques (pp. 259-277). Springer, Berlin, Heidelberg.

[13] Lee, Y. H., Wong, D. T., & Saliva: An emerging biofluid for early detection of diseases. (2009). American Journal of Dentistry, 22(4), 241-248.

[14] Rohleder, D., von Ribbeck, H. G., Brenner, T., Giessen, H., & Lemmer, U. (2012). Noninvasive continuous glucose monitoring by photoacoustic spectroscopy. Analytical Chemistry, 84(16), 6558-6565.

[15] Lukaski, H. C. (1987). Methods for the assessment of human body composition: traditional and new. American Journal of Clinical Nutrition, 46(4), 537-556.

[16] Gusev, M., Poposka, L., Spasevski, G., Kostoska, M., Koteska, B., Simjanoska, M., Ackovska, N., Stojmenski, A., Tasic, J., Trontelj, J. 2020. Noninvasive Glucose Measurement Using Machine Learning and Neural Network Methods and Correlation with Heart Rate Variability. Journal of Sensors. https://doi.org/10.1155/2020/9628281

[16] Kossowski, T., Stasinski, R. 2016. Robust IR attenuation measurement for non-invasive glucose level analysis. International Conference on Systems, Signals, and Image 75 Processing : International Conference on Systems, Signals, and Image Processing (Vol. 2016-June). https://doi.org/10.1109/IWSSIP.2016.7502770

[17] Sharma, N. K., Singh, S. 2012. Designing a non invasive blood glucose measurement sensor. 2012 IEEE 7th International Conference on Industrial and Information Systems, ICIIS 2012 : 2012 IEEE 7th International Conference on Industrial and Information Systems, ICIIS 2012. https://doi.org/10.1109/ICIInfS.2012.6304818

[18] Zhao, X., Zheng, Q., Yang, Z. M. 2016. Two types of photonic crystals applied to glucose sensor. . https://doi.org/10.1109/inec.2016.7589369

[19] Tanaka, Y., Purtill, C., Tajima, T., Seyama, M., Koizumi, H. 2017. Sensitivity improvement on CW dual-wavelength photoacoustic spectroscopy using acoustic resonant mode for 79 noninvasive glucose monitor. Proceedings of IEEE Sensors : Proceedings of IEEE Sensors. https://doi.org/10.1109/ICSENS.2016.7808685

[20] Arikan , K., Burhan, H., Sahin , E., Sen, F., A sensitive, fast, selective, and reusable enzyme-free glucose sensor based on monodisperse AuNi alloy nanoparticles on activated carbon support, Chemosphere, 291, 3 2022 , https://doi.org/10.1016/j.chemosphere.2021.132718

[21] Karimi, F., Zare, N., Bekmezci, M., Akin, M., Bayat, R., Seyitoğlu, B., Arikan, K., Isik, I., Sen, F., Enzyme-free glucose detection via scalable and economical fabrication of nickel-polyvinylpyrrolidone-modified multi-walled carbon nanotubes, Elektrochimica Acta , 496 , 8 2024 https://doi.org/10.1016/j.electacta.2024.144341

[22] Allen, J. (2007). Photoplethysmography and its application in clinical physiological measurement. Physiological Measurement, 28(3), R1-R39

[23] C. Salamea, E. Narvaez and M. Montalvo, "Database Proposal for Correlation of Glucose and Photoplethysmography Signals", Advances in Intelligent Systems and Computing, pp. 44-53, 2019. , Available: 10.1007/978-3-030-32033-1_5.

[24] Online: "Utilizing the PPG/BVP signal", support.empatica.com, 2021, Available: https://support.empatica.com/hc/en-us/articles/204954639-Utilizing-the-PPGBVP-signal

[25] D. Makowski, T. Pham, ZJ. Lau, JC. Brammer JC, F. Lespinasse, H . Pham, C. Scholzel, SHA. Chen. NeuroKit2 A Python toolbox for neurophysiological signal processing. Behavior 2021

DOI: 10.56171/ojn.1473276

Open Journal of Nano
ISSN: 2147-0081
(2024) 9–1
Research Article / Araştırma Makalesi

[26] M. Elgendi, 2012. On the Analysis of Fingertip Photoplethysmogram Signals, Current Cardiology Reviews, 8, 14-25, Available: 10.2174/157340312801215782

[27] M. Barandas, D. Folgado, L. Fernandes, S. Santos, M. Abreu, P. Bota, H. Liu, T. Schultz, H. Gamboa. TSFEL: Time Series Feature Extraction Library, SoftwareX 11 (2020). https://doi.org/10.1016/j.softx.2020.100456.

[28] Fraunhofer Portugal AICOS. (n.d.). TSFEL: Time Series Feature Extraction Library. GitHub Repository. Retrieved from https://github.com/fraunhoferportugal/tsfel

[29] DeCarlo, L. T. (1997). On the meaning and use of kurtosis and skewness. Psychological Methods, 2(3), 292-307.

[30] R. M. Kimmel, S. I. Rubin. (1997). Crest Factor. CRC Press.

[31] Ali, M. (2020). PyCaret: An open-source, low-code machine learning library in Python. Available at: https://pycaret.org.

[32] T. Nguyen, N. Tran, B. M.Nguyen, G. Nguyen, 2018. A resource usage prediction system using functional-link and genetic algorithm neural network for multivariate cloud metrics. In 2018 IEEE 11th conference on service-oriented computing and applications (SOCA), 49–56. IEEE,

[33] James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning: with Applications in R. Springer Science & Business Media.

[34] Root Mean Square Value A Dictionary of Physics (6 ed.). Oxford University Press. 2009. ISBN 9780199233991

[35] Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2004). Applied Linear Statistical Models. McGraw-Hill/Irwin.

[36] Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794.

[37] Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. International Journal of Forecasting, 22(4), 679-688.

[38] Anna Veronika Dorogush, Andrey Gulin, Gleb Gusev, Nikita Kazeev, Liudmila Ostroumova Prokhorenkova, Aleksandr Vorobev "Fighting biases with dynamic boosting". arXiv:1706.09516, 2017.

[39] Qi Meng, Guolin Ke, Taifeng Wang, Wei Chen, Qiwei Ye, Zhi-Ming Ma, Tie-Yan Liu. "A Communication-Efficient Parallel Algorithm for Decision Tree". Advances in Neural Information Processing Systems 29 (NIPS 2016), pp. 1279-1287.

[40] Moore, D. S., McCabe, G. P., & Craig, B. A. (2019). Introduction to the Practice of Statistics. W. H. Freeman.

[41] G. H Huzooree, K.K. Khedo, N. Joons, 2017. Glucose prediction data analytics for diabetic patients monitoring, 2017 1st International Conference on Next Generation Computing Applications (NextComp) 10.1109/NEXTCOMP.2017.8016197

[42] G. Edinson, G. Go´mez , L. E. B. Agredo, V. Martı´nez, O. F. B. Leiva. Application of Artificial Intelligence Techniques for the Estimation of Basal Insulin in Patients with Type I Diabetes, Research Article Open Access https://doi.org/10.1155/2020/7326073

[43] N.D Nanayakkara, S.C Munasingha, G.P. Ruwanpathirana, 2018. Non-Invasive Blood Glucose Monitoring using a Hybrid Technique, In Proceedings of the 2018 Moratuwa Engineering Research Conference (MERCon), Moratuwa, Sri Lanka, pp. 7–12. , 10.1109/MERCon.2018.8421885

DOI: 10.56171/ojn.1473276

[44] R. Sun, Y.L. Duan, Y.M. Zhang, L.G. Feng, B. Ding, R.N. Yan, J.H. Ma, X.F Su. Time in Range Estimation in Patients with Type 2 Diabetes is Improved by Incorporating Fasting and Postprandial Glucose Levels, DIABETES THERAPY, Volume14Issue8Page1373-1386, DOI10.1007/s13300-023-01432-2

[45] S. S. Gupta, S. Hossain, C. A. Haque and K. -D. Kim, 2020. In-Vivo Estimation of Glucose Level Using PPG Signal, 2020 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, Korea (South), 2020, pp. 733-736, doi: 10.1109/ICTC49870.2020.9289629

[46] Pappada, S. M., Cameron, B. D., Rosman, P. M., Bourey, R. E., Papadimos, T. J., Olorunto, W., & Borst, M. J. (2011). Neural network-based real-time prediction of glucose in patients with insulin-dependent diabetes. Diabetes technology & therapeutics, 13(2), 135-141.

[47] Aishah, A. F. Q. A., Zainuriah, M. R., & Norhilda, A. K. (2019). Multiple linear regression model analysis in predicting fasting blood glucose level in healthy subjects. In IOP Conference Series: Materials Science and Engineering (Vol. 469, No. 1, p. 012050). IOP Publishing.