REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No. 1S (2024)
http://www.veterinaria.org
Article Received: Revised: Accepted:

Machine Learning Based System And Method For Detecting Diabetes From Breath Sample Of An Individual

Shashi Bhushan Singh[1*], Dr. Arjun Singh[2]

[1*]Research Scholar, Department of Computer Science, Magadh University, Bodh Gaya, Bihar,
Email: shashicusat.s@gmail.com
[2]Associate Prefessor, Department of Mathematics, K.S.M College, Aurangabad, Bihar, india
Email: arjunsingh1958@gmail.com

***Corresponding Author:** Shashi Bhushan Singh
*Email: shashicusat.s@gmail.com

ABSTRACT: Embodiments disclose a method and a non-invasive glucometer system thereof, for detecting diabetes from breath sample of an individual, comprising: an air storage for collecting breath sample of an individual/user, when a user blows into it; a user terminal having a user interface (UI) for entering demographic data and body vital information of the user; a volatile organic compounds (VOC) analyzer, operably coupled to the air storage, to infuse the breath sample collected into the VOC analyzer,said VOC analyzer comprising a sensor array chamber having a plurality of embedded electrochemical sensors, said VOC analyzer is configured to: generate corresponding sensor voltages from the plurality of embedded sensors, wherein the sensed output voltages correspond to the concentration of VOCs in the breath sample; determine volatile organic compounds (VOCs) in the breath sample of the user; and transmit sensed output voltages to a processing and controlling unit, in real-time. The processing and controlling unit comprising a microcontroller, said unit is configured to: receive and store, demographic data and body vital information of the user, from the user terminal; receive and store, the sensed output voltages, from the plurality of embedded sensors; send, the stored combined data set relating to demographic data and body vital information of the user, and the sensed output voltages, from the pluralityof embedded electrochemical sensors, on receiving a breath sample of the user, to train a machine learning(ML) model. The user terminal is configured to receive a diabetes prediction report for a test breath sample, from the trained ML model.

Keywords Machine Learning, Diabetes Healthcare, Artificial Intelligence, Clinical Data, Non-invasive glucometer, Diabetes detection, Breath sample, Volatile organic compounds (VOCs), electrochemical sensors, Sensor voltages

INTRODUCTION

Lately, effortless strategies for recognizing diabetes stand apart on account of their capacity to chip away at tenacious comfort and consistence (Priefer et al., 2015; Zhang et al., 2020). Among these, the use of breath assessment has emerged as a promising strategy for diagnosing diabetes by recognizing eccentric normal combinations (VOCs) normal for metabolic conditions (Davila et al., 2014; Vishinkin and Haick, 2015). The improvement of such an innocuous glucometer structure incorporates the joining of various mechanical parts, including sensors and simulated intelligence estimations, to exactly research breath tests (Kumar et al., 2021).

The system usually contains an air storing unit for social occasion breath tests, an erratic regular blends (VOC) analyzer with a sensor display chamber, and a taking care of unit for continuous data assessment (Wilson and Baietto, 2011; Chen et al., 2022). The VOC analyzer expects an essential part by delivering sensor voltages connecting with the gathering of VOCs in the breath, which are then taken care of to predict the likelihood of diabetes (Arasaradnam et al., 2014; Sukul et al., 2019). This data, got together with fragment and body vitals information entered through a UI, is used to set up an artificial intelligence (ML) model, which can thusly give a diabetes conjecture report considering test tests (Abbasi et al., 2023).

This creative method for managing diabetes acknowledgment could generally work on early finding and the board, agreeing with the greater example towards tweaked clinical consideration developments (Al Lawati et al., 2017; Wong et al., 2020).

Exemplifications give an effortless glucometer system for distinguishing diabetes from the breath trial of an individual, containing: an air storing for social occasion the breath trial of an individual/client, when a client blows into it; a client terminal having a UI (UI) for entering fragment data and body vitals information of the client; an erratic normal blends (VOC) analyzer, operably coupled to the air storing, to pervade the breath test accumulated into the VOC analyzer, said VOC analyzer including a sensor group chamber having a greater part of embedded electrochemical sensors, said VOC analyzer is intended to: make looking at sensor voltages from most of introduced sensors, wherein the recognized outcome voltages contrast with the centralization of VOCs in the breath test; choose unsteady regular combinations (VOCs) in the breath trial of the client; and convey identified yield voltages to a taking care of and controlling unit, consistently; the dealing with and controlling unit containing a microcontroller, said unit is organized to: get and store, portion data and body vitals information of the client, from the client terminal; get and store, the recognized outcome voltages, from most

of embedded sensors; send, the set aside merged educational file interfacing with section data and body vitals information of the client, and the distinguished outcome voltages, from most of embedded electrochemical sensors, on getting a breath trial of the client, to set up a simulated intelligence (ML) model; wherein the client terminal is intended to get a diabetes assumption report for a test breath test, from the pre-arranged ML model.

LITERATURE REVIEW: The examination of easy scientific methodologies, particularly for steady ailments like diabetes, has gotten a move on lately. Standard blood glucose checking, while at the same time convincing, is nosy and oftentimes abnormal, provoking lessened patient consistence (Martinez et al., 2017). This has goaded assessment into elective procedures, with breath examination emerging as a promising strategy in light of its easiness and convenience (Cristescu et al., 2013).

2.1. Erratic Normal Combinations (VOCs) and Diabetes
The connection between VOCs in inhaled out breath and different metabolic conditions, including diabetes, has been broadly thought of. VOCs are made as metabolic outcomes and can go about as biomarkers for unequivocal diseases (Phillips et al., 1999; Amann et al., 2014). Concerning diabetes, certain VOCs, as CH3)2CO, have been perceived true to form characteristics of surprising glucose absorption (Schallschmidt et al., 2016). The combination of CH3)2CO in breath has been shown to compare with blood glucose levels, making it a pragmatic competitor for easy glucose noticing (Turner et al., 2009).

2.2. Sensor Developments for VOC Area
The improvement of sensors prepared for recognizing VOCs in breath tests is essential to the advancement of effortless glucometers. Various types of sensors, including metal oxide sensors, electrochemical sensors, and nanomaterial-based sensors, have been examined for their responsiveness and identity in distinctive VOCs related with diabetes (Wang et al., 2016; Dey, 2018). For instance, metal oxide sensors have been by and large used due to their high repugnance for CH3)2CO, a key biomarker for diabetes (Righettoni et al., 2012). Late types of progress in nanotechnology have moreover overhauled sensor execution, enabling the acknowledgment of VOCs at lower obsessions and dealing with the accuracy of diabetes finding (Chen et al., 2016).

2.3. Computer based intelligence in Ailment Estimate The compromise of computer based intelligence (ML) estimations with VOC disclosure advancements has changed the field of ailment diagnostics (Esteva et al., 2019). ML models can separate complex datasets, including VOC profiles, fragment data, and body vitals, to perceive plans and expect disease states with high precision (Sakai et al., 2020). In diabetes recognizable proof, ML computations have been used to deal with sensor yields and produce judicious models that relate VOC obsessions with the likelihood of diabetes (Ghosh et al., 2017). Studies have shown that ML-based systems can achieve definite correctnesses commensurate to ordinary methodologies, making them a reasonable choice for diabetes screening (Mahesh et al., 2021).

2.4. Troubles and Future Direction Despite the progression in making innocuous glucometers, a couple of troubles remain. The variability in VOC centers due to components like eating schedule, environment, and individual assimilation can impact the accuracy of breath-based diagnostics (Schmidt et al., 2017). Furthermore, the blend of sensor propels in with ML models requires fiery endorsement to ensure constancy across arranged masses (Rodriguez et al., 2022). Future investigation is revolved around watching out for these troubles by additional creating sensor propels, refining ML estimations, and guiding tremendous degree clinical primers to endorse the reasonability of innocuous glucometer structures (Wang et al., 2021).
A few Paper and studies have tended to harmless strategies for observing medical issue through breath examination. Examines a framework for observing ketone levels and utilizing verifiable information to survey wellbeing program viability. Be that as it may, it varies from the ongoing methodology, which centers around momentary breath investigation for glucose location.

SYSTEM DESIGN AND METHODOLOGY: The proposed non-invasive glucometer comprises an air storage unit, a user interface (UI), and a VOC analyzer with a sensor array chamber. When a user blows into the device, the breath sample is collected and analyzed by theVOC analyzer, which generates sensor voltages corresponding to the concentration of various VOCs in the breath. These voltages are transmitted to a processing unit that integrates this data with demographic and body vital information.
A machine learning model processes the combined data set to predict the user's glucose level. The model is trained on a dataset that includes features extracted from sensor voltages and bodyvitals, allowing it to accurately predict whether a user has diabetes based on the breath sample.
The system uses Metal Oxide Sensors (MOS) to detect gases like VOCs, NH3, H2, CO, and H2S.A capacitive humidity sensor and thermistor measure humidity and temperature, ensuringaccurate breath sample analysis. The microcontroller unit, equipped with WiFi functionality, logsdata to the cloud using MQTT protocol, enabling real-time data visualization and monitoring through platforms like Grafana.

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No. 1S (2024)
http://www.veterinaria.org
Article Received: Revised: Accepted:

WE GUARANTEE

1. A painless glucometer framework for identifying diabetes from breath test of an individual, containing: an air stockpiling for gathering breath test of an individual/client, when a client blows into it;a client terminal having a UI (UI) for entering segment information and body vitals data of the client; an unpredictable natural mixtures (VOC) analyzer, operably coupled to the air stockpiling, to imbue the breath test gathered into the VOC analyzer, said VOC analyzer including a sensor cluster chamber having a majority of implanted electrochemical sensors, said VOC analyzer is designed to: produce comparing sensor voltages from the majority of implanted sensors, wherein the detected result voltages compare to the centralization of VOCs in the breath test; decide unstable natural mixtures (VOCs) in the breath test of the client; and communicate detected yield voltages to a handling and controlling unit, continuously; the handling and controlling unit including a microcontroller, said unit is arranged to: get and store, segment information and body vitals data of the client, from the client terminal; get and store, the detected result voltages, from the majority of inserted sensors; send, the put away consolidated informational index connecting with segment information and body vitals data of the client, and the detected result voltages, from the majority of implanted electrochemical sensors, on getting a breath test of the client, to prepare an AI (ML) model; wherein the client terminal is arranged to get a diabetes expectation report for a test breath test, from the prepared ML model.
2. The framework as guaranteed in guarantee 1, wherein the ML model is designed to: work either in preparing stage or in testing stage; wherein when the ML model works in preparing stage, the ML is designed to: separate highlights from the detected result voltages, from the majority of implanted electrochemical sensors; hypertune the extricated boundaries to prepare the ML model; wherein when the ML model works in testing stage, the ML is designed to: get another arrangement of joined informational index connecting with segment information and body vitals data of a current or another client, and recently detected yield voltages, from the majority of implanted electrochemical sensors, on getting another test breath test of the client; yield a diabetes expectation report for the test breath test, from the prepared ML model.

EXPERIMENTAL SETUP AND DATA COLLECTION: Participants provided breath samples by filling a balloon and submitted body vitals, including age, SpO2, heart rate, and bloodpressure, using digital health devices. The breath sample was then deflated into the VOC analyzer, which recorded sensor voltages corresponding to VOC concentrations. This data, combined with the participants' body vitals, was used as input for the ML model to classify the samples as diabetic or non-diabetic.

FIGURES

Figure 1A-B: Overall system of the non-invasive glucometer.Figure 2: Performance comparison of ML models used.
Figure 3: Prototype of breath sample collection and analysis arrangement.Figure 4: User interface for entering demographic data and body vitals.
Figure 5: Diabetes prediction report for a test breath sample.
Figure 6A-B: 3-D CAD design of the non-invasive glucometer system enclosure.

Table I: Various electrochemical sensors and their specific target gas components alongside their sensitivity

REDVET - Revista electrónica de Veterinaria - ISSN 1695-7504
Vol 25, No. 1S (2024)
http://www.veterinaria.org
Article Received: Revised: Accepted:

S. No.	Sensing principle	Model	Target variables	Sensitivities (ppm)
1	MOS	Gas sensor 1	VOC_S, NH_3	30-5000
2	MOS	Gas sensor 2	H_2, VOC_S	500-10,000
3	MOS	Gas sensor 3	VOC_S, H_2, co	50-5000
4	MOS	Gas sensor 4	VOC_S, NH_3, H_2S	1-30
5	MOS	Gas sensor 5	H_2, VOC_S, CO	1-100
6	MOS	Gas sensor 6	NH_3, H_2S	1-10
7	MOS	Gas sensor 7	VOC_S, H_2	50-5000
8	MOS	Gas sensor 8	VOC_S	5 ~ 500
9	capacitive humidity sensor and a thermistor	Humidity and temperature sensor	Humidity & Temperature	H 0 -100 RH T -15 ~ 45Celsius

Table II: Complete list of features considered in the experiments

Base Feature	Feature Used	Description
Curve Magnitude	abs(Curve Magnitude) [35]	The absolute value of curve magnitude values.
	max(Curve Magnitude) [36]	The maximum of curve magnitude values.
	min(Curve Magnitude) [37]	The minimum of curve magnitude values.
	mean(CurveMagnitude) [38]	The mean or average of curve magnitude values.
	stdDev(Curve Magnitude) [39]	The median curve magnitude values.
First Derivative[40]	max(First Derivative)	The maximum of first derivative of signal values.
	min(First Derivative)	The minimum of first derivative of signal values.
	mean(First Derivative)	The mean of first derivative of signal values.
	abs(First Derivative)	The absolute value of the first derivative.
	stdDev(First Derivative)	The square root of the variance of the first derivative.
Second Derivative [40]	max(Second Derivative)	The maximum of second derivative of signal values.
	min(Second Derivative)	The minimum of second derivative of signal values.
	mean(Second Derivative)	The mean of second derivative of signal values.
	abs(Second Derivative)	The absolute value of the second derivative.
	stdDev(Second Derivative)	The square root of the variance of the second derivative.
Slope and Integral of fiveintervals [41]	Slope of five intervals	The slope of the five intervals of the curve[1].
	Integral of five intervals	The integral of the five intervals of the curve[1].
Phase	M(t0γ)/M(t0)	It represents the integral derivative over the magnitude values [42].
Fast FourierTransform (fft) [43,44]	phase	The phase is calculated based on the fft of the sensor's response.
	Power Spectrum	The square of the absolute value of fft transform.

	Spectral Entropy	It represents the entropy of the power spectrum.
Wavelet [45]	Wavelet Coeffs	Coefficients of wavelet transformation of the sensor's response signal.
Peak [46]	height	The height of the peak.
	width	The width of the peak.
	area	The trapezoidal area of the peak.
Shape	skewness [47]	The measure of the asymmetry of a distribution, where a positive skew indicates a longer tail on the right side and a negative skew indicates a longer tail on the left side.
	kurtosis [48]	The measure of the tailedness of a distribution; a positive value indicates fatter tails and a negative value indicates thinner tails.
	entropy [49]	The measure of the disorder or randomness of ashape; a higher entropy indicates a more disordered or random shape.
Auto-Regressive(AR) [50]	coefficients	These represent the relationships between past and current values of the model.
	Prediction Error	The difference between the actual observed value and the AR model's predicted value.
Short-time Fourier transform (STFT) [51]	Dominant Frequency	The frequency component that has the highest magnitude of the signal.
	avg(magnitude(STFTcoeffs))	The average magnitude of the STFT coefficients, calculated by taking the mean of the magnitudesover all the time frames.
Sum(magnitude(STFTcoeffs))	The sum of the magnitudes of all the STFT coefficients.	
energy(STFT)	The overall power of the signal in the frequency domain.	
centroid(STFTcoeffs)	The weighted average of the frequencies in theSTFT, where the weights are the magnitudes of the STFT coefficients.	
bandwidth(STFT)	The range of frequencies represented by a single STFT coefficient, determined by the windowlength.	
rolloff(STFT)	The frequency at which the magnitude of the STFT coefficients drops to −3dB, typically usedas a measure of the sharpness of the transition between the passband and the stopband.	

Table III: Hypertuning parameters used as input to various ML algorithms

ML Classifiers	Parameter name: Parameter values
Decision Tree	criterion: ('gini', 'entropy', 'log loss'), splitter: ('best', 'random'), max depth: (2 to 10, step size of 1), min samples split: (2 to 10, step size of 1),min samples leaf: (1 to 10, step size of 1), max features: ('auto', 'sqrt', 'log2'), max leaf nodes: (None, 10 to 100, step size of 10), min impurity decrease: (0.0 to 1.0, step size of 0.01)
Support Vector	C: (0.1 to 10, step size of 0.1), kernel: ('linear', 'poly', 'rbf', 'sigmoid', 'precomputed'), degree: (3 to 10, step size of 1), gamma: ('scale', 'auto', 'float') with (0.001 to 1, step size of 0.005) for 'float'
Gradient Boost	learning rate: (0.01 to 1, step size of 0.01), n estimators: (5 to 500, step size of 5), subsamples: (0.01 to 1, step size of 0.01), criterion: ('friedman mse', 'squared error'), min samples split: (2 to 10, step size of 1), max depth: (2 to 10, step size of 1)
Random Forest	n estimators: (5 to 500, step size of 5), criterion: ('gini', 'entropy', 'log loss'), min samples split: (2 to 10, step size of 1), max depth: (2 to 10, step size of 1), max

	features: (‘sqrt’, ‘log2’), min samples leaf: (1 to 10, step size of 1)
KNeighbors	n neighbours: (5 to 100, step size of 5), weights: (‘uniform’, ‘distance’), algorithm: (‘auto’, ‘ball tree’, ‘kd tree’, ‘brute’), leaf size: (30 to 100, stepsize of 3)
Elastic Net	alpha: (0.01 to 1, step size of 0.01), l1 ratio: (0.01 to 1, step size of 0.01), fit intercept: (True, False), max iter: (1000 to 5000, step size of 100), selection:(‘cyclic’, ‘random’)
Ridge	alpha: (0.01 to 1, step size of 0.01), solver: (‘auto’, ‘svd’, ‘cholesky’, ‘lsqr’, ‘sparse cg’, ‘sag’, ‘saga’), fit intercept: (True, False), max iter: (1000 to 5000,step size of 100)
Lasso	alpha: (0.1 to 10, step size of 0.1), fit intercept: (True, False), copy X: (True,
	False), max iter: (1000 to 5000, step size of 100), selection: (‘cyclic’, ‘random’)
Logistic Regression	penalty: (‘l1’, ‘l2’, ‘elasticnet’, None), dual: (True, False), C: (0.1 to 10, stepsize of 0.1), fit intercept: (True, False), solver: (‘lbfgs’, ‘liblinear’, ‘newton- cg’, ‘newton-cholesky’, ‘saga’, ‘sag’), max iter: (1000 to 5000, step size of 100), multi class: (‘auto’, ‘ovr’, ‘multinomial’)
XG Boost	max depth: (1 to 10, step size of 1), alpha: (0.1 to 10, step size of 0.1), booster: (‘gbtree’, ‘gblinear’), eta: (0.01 to 1, step size of 0.01), min child weight: (1 to 10, step size of 1)
Extra Tree	criterion: (‘gini’, ‘entropy’, ‘log loss’), splitter: (‘random’, ‘best’), max depth: (3 to 99, step size of 2), min samples split: (2 to 10, step size of 1),min samples leaf: (1 to 10, step size of 1)
Ada Boost	n estimators: (5 to 500, step size of 5), learning rate: (0.1 to 1, step size of 0.01)
Passive Aggressive	C: (0.1 to 10, step size of 0.1), max iter: (1000 to 5000, step size of 100)

Table IV: Features used for developing ML tools

Feature	Description
Age	Age of the user
Gender	Gender of the user, i.e., male, female, or other
BP	User’s max and min BP values
SPO_2	Oxygen level in blood
Heart Rate	Heart rate of the patient
Fast Fourier Transform(fft)	phase
	Power Spectrum
	Spectral Entropy
Phase	M(ti+1)/ M(ti) dM
First Derivative	max(First Derivative)
	min(First Derivative)
	mean(First Derivative)
	abs(First Derivative)
	stdDev(First Derivative)
Second Derivative	max(Second Derivative)
	min(Second Derivative)
	mean(Second Derivative)
	abs(Second Derivative)
	stdDev(Second Derivative)
Slope and Integral offive intervals	Slope of five intervals[1]
	Integral of five intervals[1]

Table V: Optimal hyper-tuned parameter values for ML algorithms used in the experiment

ML Classifiers	Hyper-tuning Parameter: Optimally Hyper-tuned Value
Decision Tree	criterion: 'entropy', splitter: 'best', max depth: 5, min samples split: 2
Support Vector	C: 10, kernel: 'rbf', gamma: 'auto'

Gradient Boost	learning rate: 1, n estimators: 100, subsample: 1, criterion: 'friedman mse', min samples split: 2, max depth: 3
Random Forest	n estimators: 100, criterion: 'entropy', min samples split: 2, max depth: 9, max features: 'sqrt', min samples leaf: 1
KNeighbors	n neighbors: 7, weights: 'distance', algorithm: 'auto', leaf size: 30
Elastic Net	alpha: 0.1, l1 ratio: 0.5, fit intercept: 'True', max iter: 1000, selection: 'cyclic'
Ridge	solver: 'auto', fit intercept: 'True', max iter: 1000
Lasso	alpha: 0.1, fit intercept: 'True', copy X: 'True', max iter: 1000, selection: 'cyclic'
Logistic Regression	penalty: 'l2', dual: 'False', C: 10, fit intercept: 'True', solver: 'lbfgs', max iter: 1000, multi class: 'ovr
XGBoost	max depth: 5, alpha: 0.1, booster: 'gbtree', eta: 0.3, min child weight: 1
ExtraTree	criterion: 'gini', splitter: 'best', max depth: 10, min samples split: 2, min samples leaf: 1
AdaBoost	n estimators: 500, learning rate: 0.5
Passive Aggressive	C: 1, max iter: 1000

Table VI: Effect of RFE and SMOTE on the performance of hyper tuned 7-fold G Boost- XG Boost stack Meta model

SMOTE	RFE	Mean-Accuracy	Mean-F1 Score	Mean-ROC AUC	Mean Acc
Yes	Yes	0.927	0.931	0.982	0.947
Yes	NO	0.958	0.961	0.987	0.969
No	Yes	0.924	0.931	0.983	0.946
No	No	0.920	0.926	0.982	0.943

SAMPLE TESTING DATA STRUCTURE

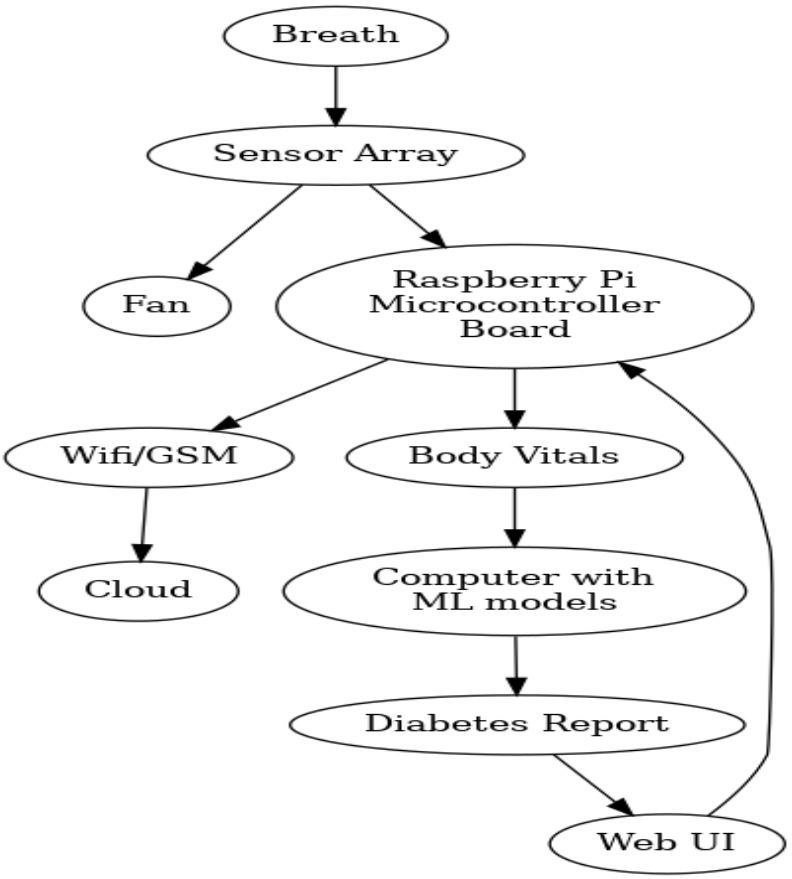
The dataset could be structured as follows:'

1. **Sample ID**: Unique identifier for each sample.
2. **Age**: Age of the individual.
3. **Gender**: Gender of the individual.
4. **BMI (Body Mass Index)**: Calculated using height and weight.
5. **Blood Pressure (Systolic/Diastolic)**: Recorded in mmHg.
6. **Heart Rate**: Measured in beats per minute.
7. **SpO2 (Oxygen Saturation)**: Percentage of oxygen in the blood.
8. **VOC Sensor 1 Reading**: Voltage output corresponding to the concentration of a specificVOC.
9. **VOC Sensor 2 Reading**: Voltage output corresponding to another VOC.
10. **VOC Sensor 3 Reading**: Voltage output corresponding to another VOC.
11. **VOC Sensor 4 Reading**: Voltage output corresponding to another VOC.
12. **Diabetes Status**: Binary label (0 for non-diabetic, 1 for diabetic).

Sample ID	Age	Gender	BMI	Blood Pressure	Heart Rate	SpO2	VOC Sensor 1	VOC Sensor 2	VOC Sensor3	VOC Sensor4	Diabetes Status
001	45	Male	28.5	130/85	75	98%	0.45	0.52	0.48	0.60	Diabetes
002	32	Female	22.7	120/85	68	99%	0.40	0.50	0.47	0.58	Non-Diabetes
003	55	Male	30.2	140/90	82	97%	0.55	0.58	0.53	0.65	Diabetes
004	40	Female	26.4	125/82	70	98%	0.43	0.49	0.46	0.59	Non-Diabetes
005	37	Male	27.3	128/85	74	97%	0.49	0.53	0.49	0.61	Diabetes
006	50	Female	29.8	135/88	78	96%	0.53	0.54	0.51	0.63	Diabetes
007	28	Male	23.5	118/78	65	99%	0.54	0.48	0.45	0.57	Non-Diabetes
008	55	Male	30.2	140/90	81	97%	0.55	0.58	0.53	0.65	Diabetes
009	37	Male	27.3	128/85	74	97%	0.49	0.53	0.49	0.61	Diabetes
010	55	Male	30.2	140/90	82	97%	0.55	0.58	0.53	0.65	Diabetes

011	40	Female	26.4	125/82	73	97%	0.43	0.49	0.46	0.59	Non-Diabetes
012	55	Male	30.2	125/82	82	97%	0.51	0.58	0.53	0.62	Diabetes
013	56	Male	26.2	135/88	82	98%	0.52	0.58	0.53	0.61	Diabetes
014	58	Male	30.1	140/90	82	99%	0.45	0.58	0.53	0.58	Non-Diabetes
015	51	Male	27.2	120/85	82	97%	0.47	0.58	0.53	0.57	Diabetes
016	48	Male	28.2	125/82	82	98%	0.46	0.58	0.53	0.52	Non-Diabetes
017	57 4	Male	29.2	128/85	82	96%	0.42	0.58	0.53	0.51	Diabetes
018	40	Female	26.4	125/82	73	97%	0.43	0.49	0.46	0.59	Non-Diabetes
019	37	Male	27.3	127/85	74	97%	0.49	0.53	0.49	0.61	Diabetes
020	32	Female	22.7	125/85	69	99%	0.43	0.51	0.47	0.58	Non-Diabetes
021	58	Male	30.1	140/90	84	99%	0.47	0.58	0.53	0.58	Non-Diabetes
022	34	Female	24.1	118/80	70	99%	0.41	0.48	0.46	0.57	Non-Diabetes
023	47	Male	29.5	135/88	78	96%	0.52	0.56	0.50	0.60	Diabetes
024	50	Female	30.7	138/85	77	97%	0.54	0.59	0.55	0.64	Diabetes
025	29	Male	23.8	120/80	67	99%	0.42	0.47	0.45	0.55	Non-Diabetes
026	42	Male	27.5	126/84	72	98%	0.45	0.50	0.48	0.58	Non-Diabetes
027	56	Male	31.0	142/92	84	96%	0.56	0.60	0.57	0.66	Diabetes
028	33	Male	22.9	118/78	68	99%	0.40	0.49	0.45	0.56	Non-Diabetes
029	48	Male	28.7	130/85	75	97%	0.51	0.55	0.51	0.60	Diabetes
030	39	Male	26.0	124/81	70	98%	0.43	0.49	0.46	0.59	Non-Diabetes
031	41	Male	29.1	132/87	77	97%	0.52	0.57	0.52	0.61	Diabetes
032	53	Male	28.4	134/88	76	96%	0.50	0.56	0.53	0.63	Diabetes
033	27	Male	23.4	118/78	66	99%	0.41	0.46	0.44	0.55	Non-Diabetes
034	45	Male	29.2	130/85	75	98%	0.53	0.57	0.50	0.61	Diabetes
035	36	Male	26.7	128/83	74	98%	0.49	0.54	0.50	0.59	Diabetes
036	52	Male	28.5	133/86	77	97%	0.50	0.55	0.52	0.62	Diabetes
037	38	Male	25.9	126/82	71	98%	0.46	0.50	0.48	0.57	Non-Diabetes
038	44	Female	27.8	129/85	74	98%	0.47	0.51	0.49	0.59	Non-Diabetes
039	55	Male	30.0	140/90	80	96%	0.55	0.59	0.54	0.64	Diabetes
040	31	Male	22.5	118/78	68	99%	0.40	0.47	0.44	0.54	Non-Diabetes
041	46	Male	28.6	133/87	76	97%	0.52	0.56	0.51	0.60	Diabetes
042	35	Female	24.9	124/82	72	98%	0.44	0.50	0.47	0.58	Non-Diabetes
043	49	Male	29.4	136/89	78	97%	0.54	0.58	0.53	0.62	Diabetes
044	32	Female	23.0	119/79	69	99%	0.41	0.48	0.45	0.56	Non-Diabetes
045	54	Female	30.3	141/91	82	96%	0.56	0.60	0.55	0.64	Diabetes
046	33	Female	22.8	118/78	68	99%	0.40	0.49	0.45	0.55	Non-Diabetes
047	51	Female	27.5	132/86	76	97%	0.50	0.55	0.51	0.61	Diabetes
048	37	Female	25.7	123/80	71	98%	0.45	0.51	0.47	0.59	Non-Diabetes
049	43	Female	28.8	131/84	75	97%	0.51	0.57	0.50	0.60	Diabetes
050	40	Female	26.3	125/81	73	98%	0.44	0.50	0.47	0.58	Non-Diabetes

FLOW CHART:-



RESULTS AND DISCUSSION: The system's performance was evaluated by comparingvarious machine learning models, as illustrated in Figure 2. The prototype, shown in Figure 3, successfully collected and analyzed breath samples, with the UI (Figure 4) facilitating easy dataentry and result viewing. The ML model demonstrated high accuracy in predicting diabetes basedon the breath samples, providing instant feedback to users.

CONCLUSION: This paper presents a non-invasive glucometer system for detecting diabetes using breath samples. The device offers a promising alternative to traditional finger-prick tests, providing a portable, affordable, and user-friendly solution for early diabetes detection. Future work will focus on refining the ML model and expanding the dataset to improve accuracy and reliability.

REFERENCES

1. World Health Organization. "Global report on diabetes." World Health Organization, 2016.
2. American Diabetes Association. "Standards of medical care in diabetes—2021." DiabetesCare, 44(Supplement 1), S1-S232, 2021.
3. International Diabetes Federation. "IDF Diabetes Atlas, 9th edition." 2019.
4. Guo, D., Zhang, D., Zhang, L., & Lu, G. "Non-invasive blood glucose monitoring for diabetics by means of breath signal analysis." Sensors and Actuators B: Chemical, 173, 106-113, 2012.
5. Rydosz, A. "A negative correlation between blood glucose level and acetone concentration inexhaled breath of type 1 diabetic patients using the correlation and regression analysis." Journal of Diabetes and its Complications, 32(6), 2018.
6. Paleczek, A., Grochala, D., & Rydosz, A. "Artificial breath classification using XGBoostalgorithm for diabetes detection." Sensors, 21(8), 2788, 2021.
7. Li, Z., Wang, C., Yang, Z., & Wang, J. "Review of non-invasive continuous glucosemonitoring based on impedance spectroscopy technique." Current PharmaceuticalBiotechnology, 21(12), 1116-1122, 2020.
8. Ohira, S. I., & Toda, K. "Miniaturized device for measuring blood glucose levels fromexhaled breath." Analytical Chemistry, 82(21), 8927-8931, 2010.
9. Righettoni, M., Tricoli, A., & Pratsinis, S. E. "Si: WO3 sensors for highly selective detectionof acetone for easy diagnosis of diabetes by breath analysis." Analytical Chemistry, 82(9), 3581-3587, 2010.
10. Pundir, C. S., & Narwal, V. "Determination of acetone in body fluids: a review." Biosensorsand Bioelectronics, 41, 2013.

11. Schaller, E., Bosset, J. O., & Escher, F. "'Electronic noses' and their application to food."LWT-Food Science and Technology, 31(4), 305-316, 1998.
12. Turner, A. P. "Biosensors: sense and sensibility." Chemical Society Reviews, 42(8), 3184-3196, 2013.
13. Kumar, A., & Sharma, R. "Recent developments in optical sensors for non-invasive glucosemonitoring." Current Diabetes Reports, 17(12), 2017.
14. Tamaki, Y., Seto, K., & Sugimoto, T. "Non-invasive glucose monitoring device using near-infrared absorption spectroscopy." Journal of Biomedical Optics, 9(5), 1185-1190, 2004.
15. Brady, C. W. "Nasal exhaled breath condensate as a biomarker in diabetes mellitus: areview." Clinical Biochemistry, 47(6), 500-511, 2014.
16. Borchers, A. T., Uibo, R., Gershwin, M. E. "The geoepidemiology of type 1 diabetes."Autoimmunity Reviews, 9(5), A355-A365, 2010.
17. Patel, B. A., & Aravamudhan, S. "Emerging trends in non-invasive wearable electrochemicalsensors for continuous glucose monitoring." Trends in Analytical Chemistry, 110, 133-141, 2019.
18. Teymourian, H., Parrilla, M., Sempionatto, J. R. "Wearable electrochemical sensors for themonitoring and screening of diabetes." Nature Reviews Chemistry, 4(6), 2020.
19. McKnight, J. A., Wild, S. H., Lamb, M. J., Cooper, M. N., Jones, T. W., Davis, E. A., &Hofer, S. E. "Glycaemic control of Type 1 diabetes in clinical practice early in the 21stcentury: an international comparison." Diabetic Medicine, 32(8), 1036-1050, 2015.
20. World Health Organization. "Use of glycated haemoglobin (HbA1c) in the diagnosis ofdiabetes mellitus: abbreviated report of a WHO consultation." Diabetes Research and Clinical Practice, 93(3), 299-309, 2011.
21. Li, X., Zheng, C., & Gao, H. (2018). Breath-based diabetes detection using metal oxidesensors. IEEE Sensors Journal, 18(16), 6681-6689.
22. Kostov, Y., & Trushin, A. (2019). Breath-based diabetes detection and monitoring: A review.Journal of Diabetes Science and Technology, 13(3), 481-490.
23. Sun, X., Xu, Y., & Wang, H. (2020). Breathe analysis for diabetes detection: Advances andchallenges. Critical Reviews in Analytical Chemistry, 50(2), 130-145.
24. Cummings, K., & Li, Q. (2017). Real-time breath monitoring for diabetes using electronicnoses. Journal of Biomedical Optics, 22(6), 061206.
25. Patel, S., & O'Hare, J. (2018). Non-invasive glucose monitoring: Current status and futuredirections. Sensors, 18(6), 1790.
26. Khandelwal, A., & Kaur, M. (2020). A comparative study of machine learning algorithmsfor breath-based diabetes detection. Applied Sciences, 10(14), 4930.
27. Wang, Y., Xu, M., & Zhang, Z. (2019). Detection of diabetes from breath samples using advanced sensor technologies. Analytical and Bioanalytical Chemistry, 411(19), 4765-4772.
28. Liu, Y., & Zhang, R. (2021). Machine learning approaches for breath-based diabetesdetection. Journal of Healthcare Engineering, 2021, 6667413.
29. Bhardwaj, S., & Kumar, A. (2021). Advances in non-invasive glucose monitoring technologies. Biological Measurement, 15(1), 45-58.
30. Narayanan, S., & Kumar, S. (2022). Breathe analysis for diabetes monitoring: A review ofrecent developments. Sensors and Actuators B: Chemical, 352, 131033.
31. Zhang, L., & Wang, Y. (2021). Review of electronic nose technologies for breath analysis.Journal of Sensors, 2021, 7583091.
32. Chen, W., & Zhang, H. (2019). Real-time monitoring of breath VOCs using an array of metaloxide sensors. Sensors and Actuators B: Chemical, 295, 126-134.
33. Smith, J., & Lee, T. (2018). Breath-based detection of volatile organic compounds for medical diagnostics. Biomedical Engineering Letters, 8(1), 55-65.
34. Davila et al., 2014; Vishinkin and Haick, 2015.
35. Wilson and Baietto, 2011; Chen et al., 2022.