iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** XI **Month of publication:** November 2022

DOI: https://doi.org/10.22214/ijraset.2022.47262

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XI Nov 2022- Available at www.ijraset.com

Healthcare Monitoring System for Diabetic Patients Using Machine Learning

Suraka Maha Lakshmi Reddy[1], Adusumilli Yagna Gayathri[2]
[1, 2]*Department of computer science and Engineering, G. Narayanamma Institute of Technology and science for women, Telangana, India*

Abstract: ***Present-day technology provides quite an efficient way of monitoring an individual's health. Bluetooth Low Energy (BLE)-based sensors can be considered as one of the solutions for checking personal vital signs data such as blood pressure, heart rate, weight, and blood glucose (BG). In this study, we propose a personalized healthcare monitoring system by utilizing a BLE-based sensor device dataset, data processing, and machine learning-based algorithms to help diabetic patients to better self-manage their chronic condition. The proposed real-time data processing system utilizes machine Learning algorithms to train the model. Machine learning–based classification methods were tested on a diabetes dataset in order to show that a Multilayer Perceptron can provide early prediction of diabetes given the user's sensor data as input. Furthermore, the proposed diabetes classification and prediction might be integrated with individualized diet and physical activity recommendations to improve patients' health quality and avert severe circumstances in the future.***

I. INTRODUCTION

Diabetes mellitus, or polygenic disorder, has become a worldwide epidemic. polygenic disorder may be a semipermanent metabolic unwellness within which blood sugar (BG) levels fluctuate. It's caused by either skimpy internal secretion synthesis within the body, which is referred to as sort one diabetes or the body's inability to use the internal secretion it produces, that is thought as sort of a pair of polygenic disorders. sort a pair of polygenic disorders, that accounts for 90–95% of all cases of polygenic disorder and may be a growing epidemic that poses an incredible burden on health care systems, particularly in developing countries, has seen a bigger increase in diagnosing. In 2014, 4.8 million Koreans (13.7 p.c of the population) had polygenic disorder, however in 2015, 30.3 million Americans of all ages (9.4% of the population) had polygenic disorder. polygenic disorder is predicted to extend to 366 million folks worldwide by 2030, up from 171 million in 2000. Diabetic complications like upset, pressure|cardiovascular disease}, and stroke will occur if the unwellness isn't properly managed. Regular glucose observance, on the opposite hand, is vital for avoiding and decreasing diabetic complications. Recent advances in info and communication technology (ICT), still as novel biosensors that permit period of time observance of a patient's condition, have given polygenic disorder management a brand new lease on life. Diabetic patients will monitor aldohexose changes with a self-monitoring {blood aldohexose|blood sugar|glucose} (SMBG) transportable device and continuous glucose observance (CGM) sensors, and respond instantly with the mandatory action. The findings show that observance of patients' aldohexose levels will facilitate them improve malady} management and find higher management of their disease. A aldohexose observance system with sensors, an entranceway (smartphone), and a cloud system is the best choice for higher polygenic disorder management. It employs a smartphone as an entranceway to gather detector information from a body-worn detector node. The detector node's communication with the smartphone necessitates wireless technology still as low-power operation, and therefore the best resolution for this is often Bluetooth Low Energy (BLE).

The number of BLE-enabled wearable sensors, mobile devices (smartphones), and different connected things is steadily increasing. This is often clearly the time of huge information, given the immense volume of detector information collected by a detector device. There are some smart studies on however huge information analytics are often employed in health care. The NoSQL info (DB) is the greatest resolution for processing monumental amounts of detector information since it offers wonderful performance still because of the potential for health care applications. Using BLE-based sensors and period of time processing, and has a tendency to gift a tailored healthcare observance system for polygenic disorder patients during this study. The latter handles incoming detector information with Apache author, whereas the unstructured detector information is kept with MongoDB. The big quantity of continuous information (e.g., BG, heart rate, blood pressure, weight, and different personal data) from BLE-based detector devices could also be handled in a period of time by applying a period of time processing. Moreover, a Multilayer Perceptron (MLP)-based classification algorithmic program is employed to reason the diabetic patient, whereas Long memory (LSTM) is used to forecast the BG level.

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XI Nov 2022- Available at www.ijraset.com

II. OBJECTIVES

The motto of the work is to:

1) Predict the diabetic status of the patient.
2) Predict the health risk of the patient based on various characteristics.
3) Selecting Multilayer Perceptron (MLP)-based classification algorithm and Long short-term memory (LSTM) for predicting.
4) Build the model using the above-mentioned algorithms.
5) Test the model for the mentioned machine learning algorithm based on the test set.

III. METHODOLOGY

A. Machine Learning

The phrase "machine learning" refers to the discovery of important patterns in data by a computer. It's become a common technique in virtually every activity that requires data extraction from large data sets during the last few decades. We are surrounded by machine learning-based technology: search engines learn how to bring North American countries the most basic results (while inserting pro-table ads), anti-spam code learns to filter our email messages, and Mastercard transactions are protected by a code that learns how to detect frauds. Digital cameras are learning to recognize faces, while intelligent personal assistant apps on smartphones are learning to recognize spoken instructions. Accident-avoidance systems are built into automobiles using victimization machine learning techniques. In addition, machine learning is widely used in scientific fields such as bioinformatics, medicine, and physical science. One aspect of all of these applications is that, unlike other more traditional uses of computers, due to the complexity of the patterns that must be recognized, a person's applied scientist cannot provide a clear, finely detailed definition of how such activities should be completed. Several of our talents are non-inheritable or may be reined via learning from our expertise, as evidenced by intelligent creatures (rather than following specific directions given to us). Machine learning technologies are used to provide programmers with the ability to discover and adapt.

B. Dataset

A dataset of female Pima Indian patients with a minimum age of twenty-one years was pulled from the UCI machine learning repository. The National Institute of Diabetes and Digestive and Kidney Diseases previously controlled this data. There are 768 cases in this dataset, divided into two groups: There are eight risk factors that both diabetics and non-diabetics as mentioned below,

1) Pregnancies
2) Glucose
3) Blood Pressure
4) Skin Thickness
5) Insulin
6) BMI
7) Diabetes Pedigree Function
8) Age
9) Outcome

C. Preprocessing

Preprocessing stage data is used to reduce model mistakes as much as possible. Preprocessing information ensures that the model employs the most accurate dataset available because a model is only as good as the data it is fed. While one won't be able to completely clean the dataset but will be able to take a few easy actions to ensure that the data has the best chance of being used to construct an accurate model. In the dataset, search for null values as the null values in a dataset are useless and empty. Because one'll be focusing all of the efforts on "connecting" useless data rather than valuable information. If one doesn't remove null values from a model, it will incorrectly forecast results, lowering the accuracy rate.

The diabetes data set originated from UCI Machine Learning Repository.

```
 import pandas as pd
 import numpy as np
 import matplotlib.pyplot as plt                %matplotlib inline
 diabetes = pd.read_csv('diabetes.csv')
print(diabetes.columns)
```

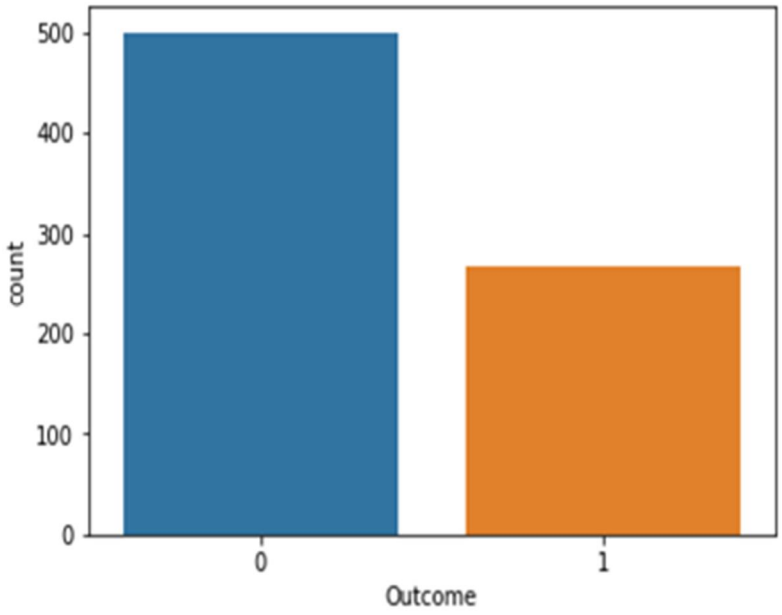
©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XI Nov 2022- Available at www.ijraset.com

Index([‘Pregnancies’, ‘Glucose’, ‘BloodPressure’, ‘SkinThickness’, ‘Insulin’, ‘BMI’, ‘DiabetesPedigreeFunction’, ‘Age’, ‘Outcome’], dtype=’object’)
diabetes.head()

	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	DiabetesPedigreeFunction	Age	Outcome
0	6	148	72	35	0	33.6	0.627	50	1
1	1	85	66	29	0	26.6	0.351	31	0
2	8	183	64	0	0	23.3	0.672	32	1
3	1	89	66	23	94	28.1	0.167	21	0
4	0	137	40	35	168	43.1	2.288	33	1

The diabetes data set consists of 768 data points, with 9 features each:
print("dimension of diabetes data: {}".format(diabetes.shape))
dimension of diabetes data: (768, 9)
“Outcome” is the feature that is going to be predicted, 0 means No diabetes, 1 means diabetes. Of these 768 data points, 500 are labeled as 0 and 268 as print(diabetes.groupby('Outcome').size())



import seaborn as sns.countplot(diabetes['Outcome'],label="Count")
diabetes.info()

```
RangeIndex: 768 entries, 0 to 767
Data columns (total 9 columns):
Pregnancies                 768 non-null int64
Glucose                     768 non-null int64
BloodPressure               768 non-null int64
SkinThickness               768 non-null int64
Insulin                     768 non-null int64
BMI                         768 non-null float64
DiabetesPedigreeFunction    768 non-null float64
Age                         768 non-null int64
Outcome                     768 non-null int64
dtypes: float64(2), int64(7)
memory usage: 54.1 KB
```

©IJRASET: All Rights are Reserved

D. Functional Requirement

1) The application takes the health-related data from the users.

2) The interface allows the users to know the probability of future risk based on the data entered in the form. The form takes 8 inputs from the user, they are

a) Number of times pregnant
b) Plasma glucose concentration 2 hours in an oral glucose tolerance test
c) Diastolic blood pressure (mm Hg)
d) Triceps skinfold thickness (mm)
e) 2-Hour serum insulin (mu U/ml)
f) Body mass index (weight in kg/(height in m)^2)
g) Diabetes pedigree function
h) Age (years)

3) The dataset is collected to perform prediction analysis.

4) Data preprocessing is used for transforming raw data into an understandable format.

5) 70% of data is divided into the training set and the remaining 30% is for the testing set. After a model has been processed by using the training set, you test the model by making predictions against the test set.

6) Multilayer Perceptron (MLP) is a deep learning method which is also a feedforward artificial neural network that generates a set of outputs from a set of inputs

7) You train the model using the training set. Test the model means to test the accuracy of the model.

8) A Classification report is used to measure the quality of predictions from a classification algorithm. The report shows the main classification metrics precision, recall, and f1-score on a per-class basis.

9) Long Short Term Memory (LSTM) is going to be used for making future predictions accurately and efficiently.

10) This application can be used by doctors and patients themselves.

IV. ANALYSIS AND DESIGN

A. Use Case Diagram

Use case diagrams represent the situation of the system. A situation is nothing, but a sequence of steps describing AN interaction between a user and a system. Thus a use case may be a set of situations tied along by some goal.

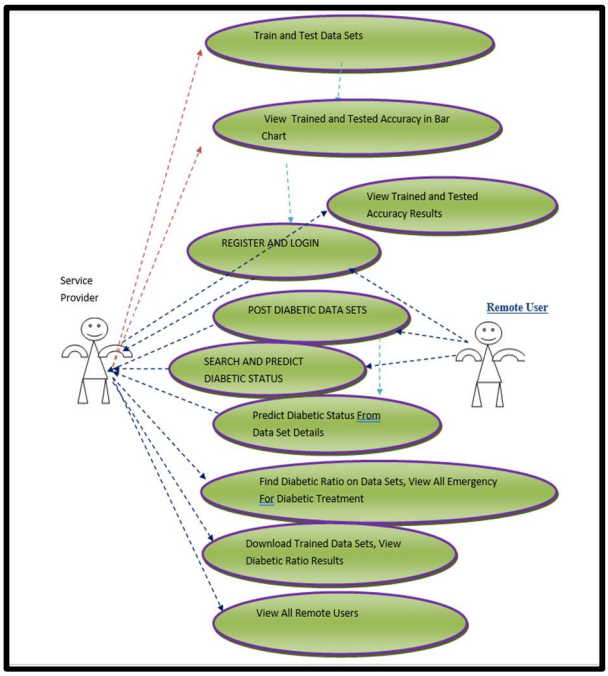


Fig 1-Use case diagram

A. Data Flow Diagram

A data flow diagram (DFD) maps out the flow of knowledge for any method or system. It uses outlined symbols like rectangles, circles and arrows, and short text labels, to indicate knowledge inputs, outputs, storage points and therefore the routes between every destination.

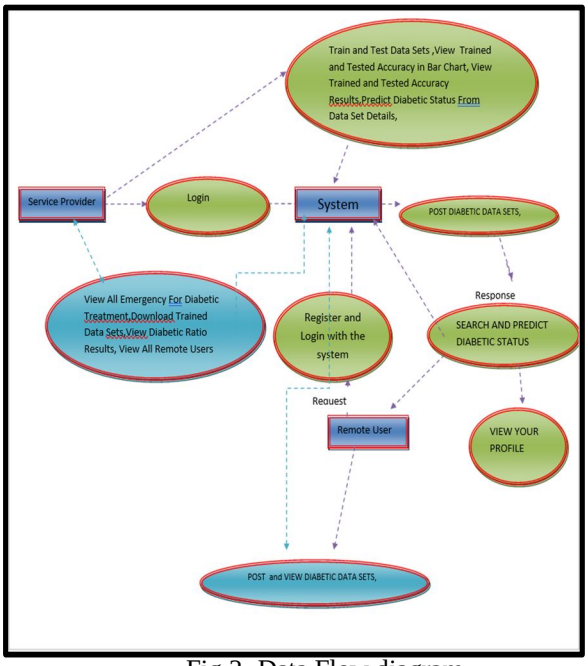


Fig 2- Data Flow diagram

B. Class Diagram

In software package engineering, a category diagram within the Unified Modeling Language (UML) may be a form of static structure diagram that describes the structure of a system by showing the system's categories, their attributes, operations (or methods), and also the relationships among objects.

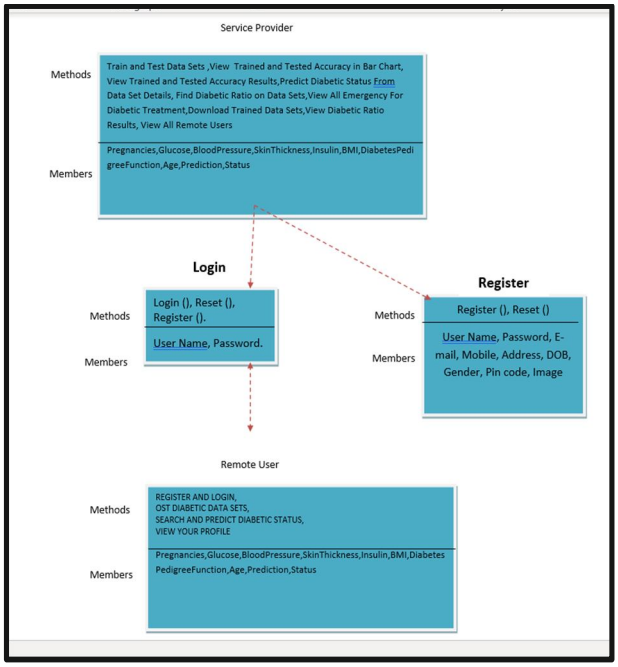


Fig 3- Class diagram

C. Sequence Diagram

In the sequence diagram however the article interacts with the opposite object is shown. There square measure sequence of events that square measure diagrammatic. It's an amount settled perspective on the association between things to attain a conduct objective of the framework

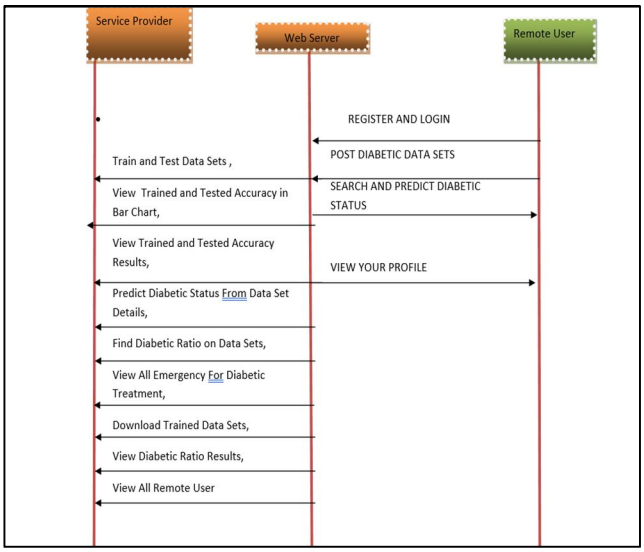
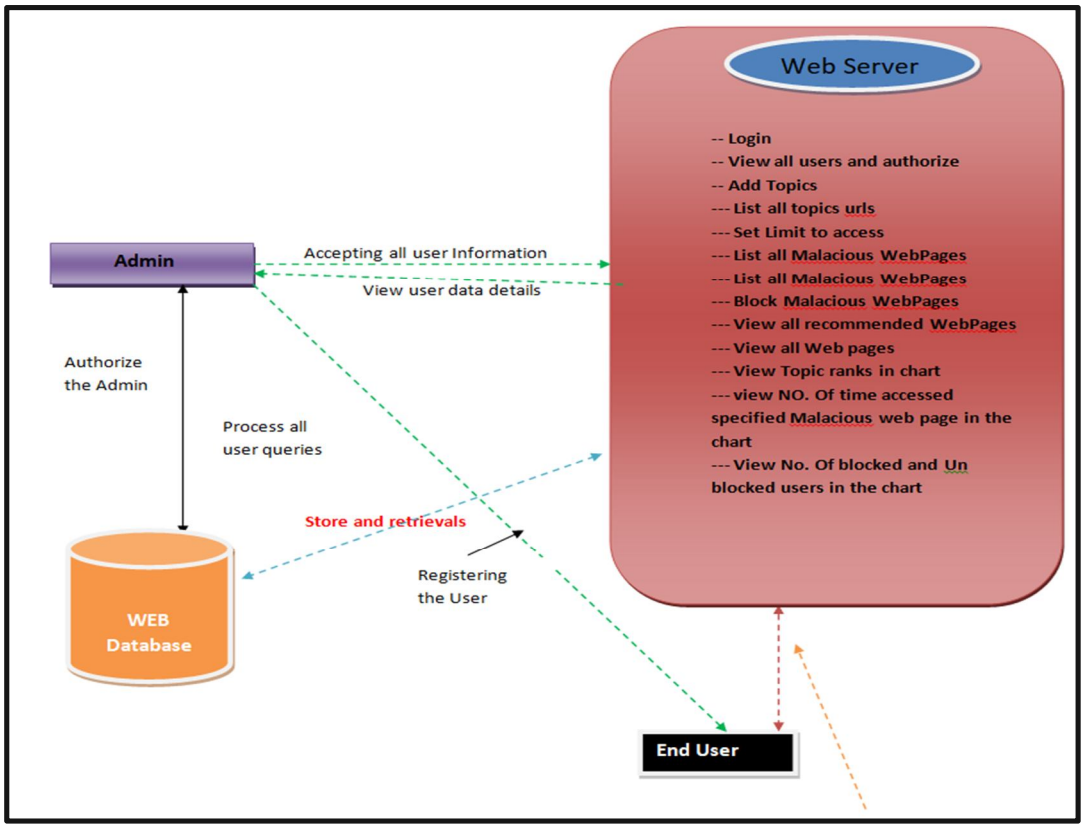


Fig 4- Sequence diagram

V. SYSTEM ARCHITECTURE



A. Flow Chart

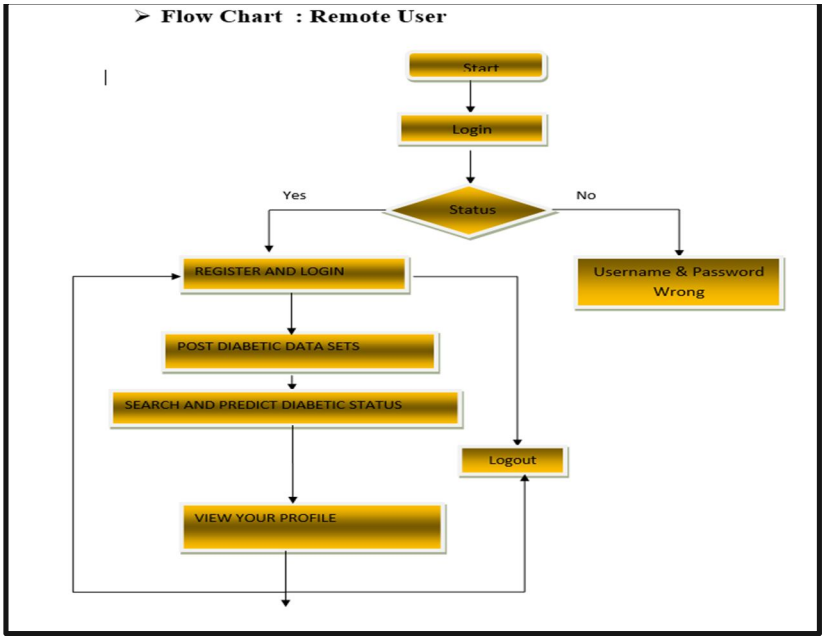


Fig 5(a)- Remote User

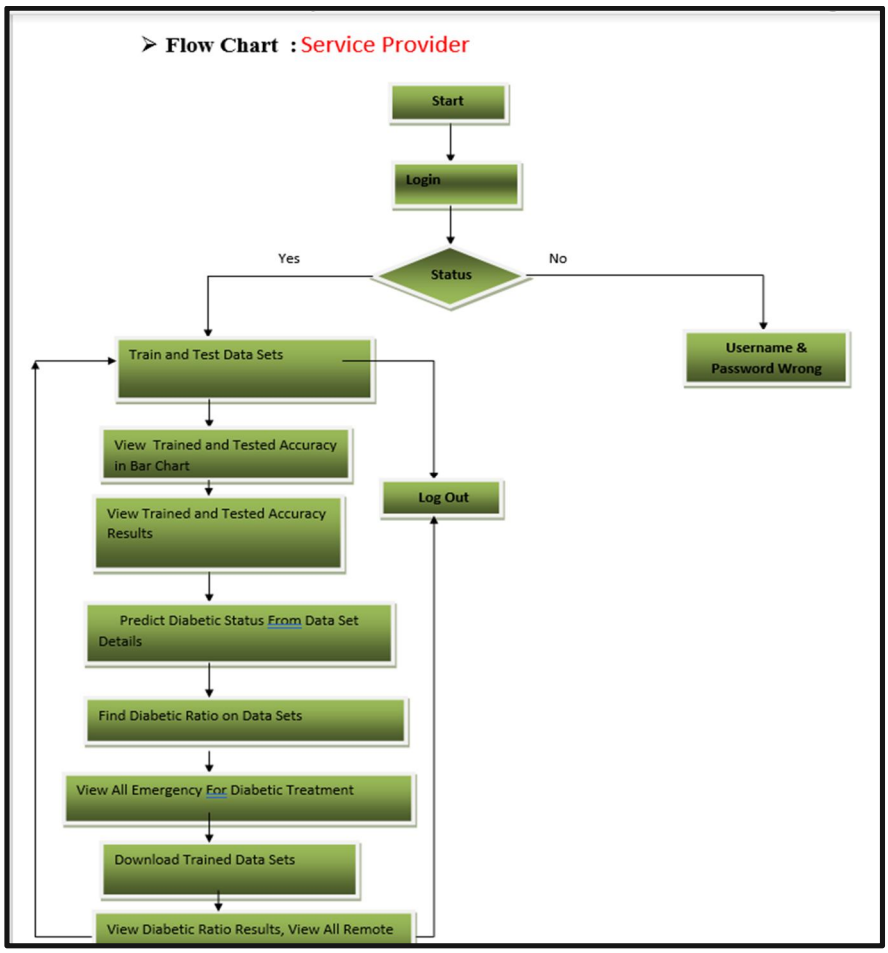


Fig 5(b) –Service Provider

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XI Nov 2022- Available at www.ijraset.com

VI. IMPLEMENTATION

The implementation of the project is done with the help of python language. To be particular, for the purpose of machine learning Anaconda is being used. The algorithms used are Multilayer perceptron (MLP) and Long short term memory (LSTM).

A. Multilayer Perceptron (MLP)

The field of artificial neural networks is commonly just cited as neural networks or multi-layer perceptrons once maybe the foremost helpful variety of neural networks. A perceptron may even be one somatic cell model that was a precursor to larger neural networks. It is a field that investigates however easy models of biological brains are about to be accustomed solve laborious machine tasks like the clairvoyant modeling tasks got AN inclination to determine in machine learning. The goal isn't to create realistic models of the brain, however instead to develop sturdy algorithms and knowledge structures that have a tendency to model laborious issues.

The power of neural networks comes from their ability to look out the illustration in your employment knowledge and so the due to best relate it to the output variable simply just have to be compelled to predict. throughout this sense, neural networks learn a mapping. Mathematically, sq. measure capable of learning any mapping perform and are tested to be a universal approximation rule.

The clairvoyant capability of neural networks comes from the hierarchic or multi-layered structure of the networks. the data structure will notice (learn to represent) selections at completely fully completely different scales or resolutions and mix them into higher-order selections.

B. Long Short Term Memory (LSTM)

Long Short Term Memory could also be a fairly recurrent neural network. In RNN output from the last step is fed as input inside this step. LSTM was designed by Hochreiter & Schmidhuber. It tackled the matter of long dependencies of RNN throughout that the RNN will not predict the word keep inside the basic cognitive process but offers tons of correct predictions from the recent data. as a result of the gap, length can increase RNN does not provide degree economic performance. LSTM can by default retain the data for associate extended quantity of some time. it's used for the method, predicting, and classifying on the premise of time-series data.

VII. DATA VISUALIZATION

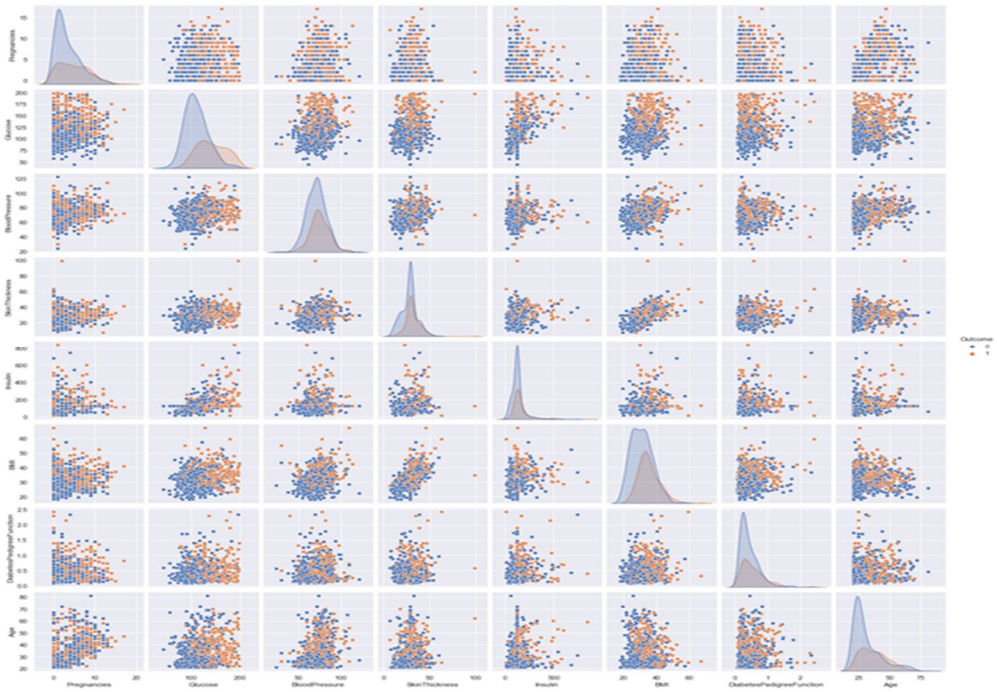


Fig 6- Pair Plot

©IJRASET: All Rights are Reserved

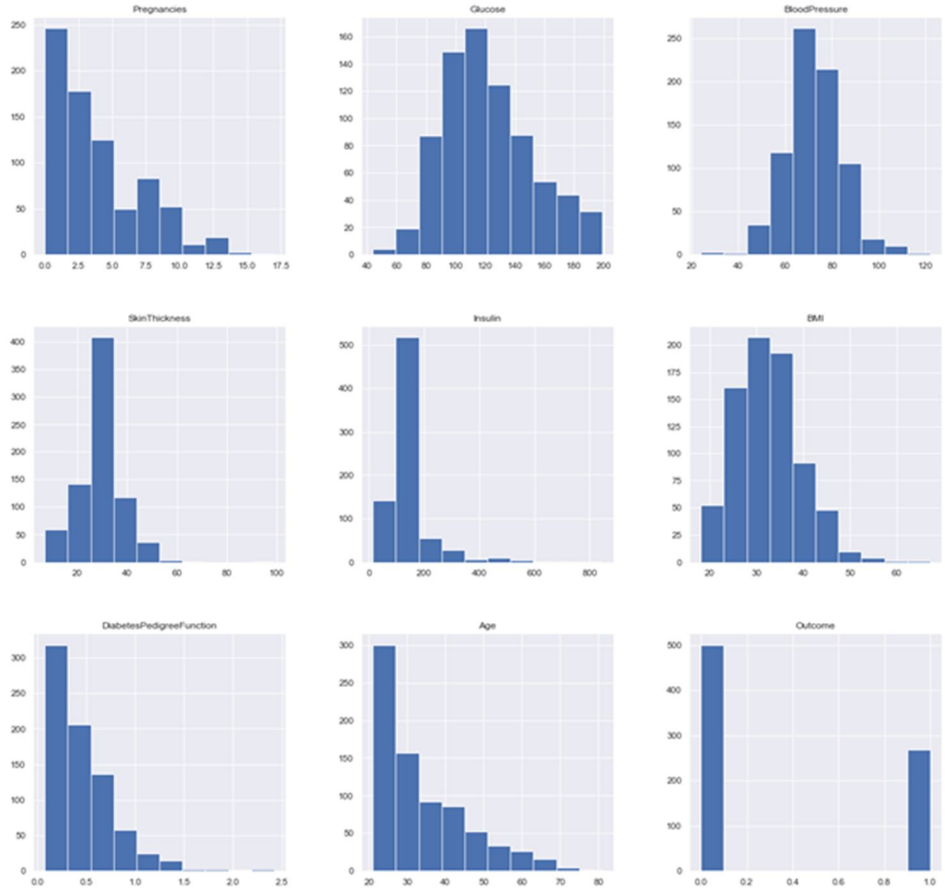


Fig 7- Histogram



Fig 8- Heat map

VIII. CONCLUSION AND FUTURE SCOPE

A. *Conclusion*

The proposed model should allow users to monitor their vital signs data from a BLE-based sensor via their smartphone. Furthermore, the suggested approach assists users in detecting the danger of diabetes at an early stage as well as assisting patients in obtaining future BG level forecasts. As a result, users can avoid the worst-case scenarios in the future. As demonstrated in this work that by combining BLE-based sensor devices with the suggested personalized healthcare monitoring system, the whole history of a user's vital signs data can be collected and evaluated.

The healthcare monitoring system should be scalable to handle the increasing volume of sensor data from BLE-based sensor devices as well as the growing number of patients. In this study, real-time data processing using Apache Kafka and NoSQL MongoDB was proposed. The size of sensor documents and the number of clients had an impact on the system's throughput and latency, according to the results.

Furthermore, when compared to the old paradigm, the proposed real-time data processing allowed the healthcare monitoring system to process a large volume of continuous user sensor data more efficiently. Furthermore, the performance of the BLE-based sensors was evaluated using a variety of metrics, including average data packet delivery and CPU/memory consumption of the android app. The distance between the BLE-based sensor and the smartphone influenced data packet delivery performance, while data transmission to the server affected CPU and memory use. However, the BLE-based sensor devices effectively captured and transferred the user's vital signs data to the smartphone within an acceptable distance in all of the trial settings. Furthermore, because the app only required low CPU and memory utilization, current commercially available smartphones are capable of serving as gateways for transmitting sensor data to the server.

Based on the input of user vital signs data collected by BLE-based sensor devices, we propose diabetes classification and BG prediction. The MLP-based classification method was used to classify diabetes patients, while LSTM was employed to predict blood glucose levels.

In comparison to the other models examined, both prediction models (MLP and LSTM) demonstrated substantial results by delivering high accuracy and low error.

The outcomes of this study are expected to be used by two groups of users: healthy people who haven't been diagnosed with diabetes yet, and diabetic patients. The proposed healthcare monitoring system can be utilized by normal people to assess their present health state and anticipate whether they will develop diabetes in the future, while diabetic sufferers can use it as a personal tool to regulate and self-manage their chronic disease. Diabetic patients can also get future BG forecasts and, by learning this information as soon as possible, can prevent their condition from worsening in the future. Because the diabetes classification dataset in this work was limited to PIMA Indian women, it is challenging to generalize the robust classification algorithm for use in other situations. The accuracy of the classification model will improve if an actual dataset is acquired from a real-world implementation.

B. *Future Scope*

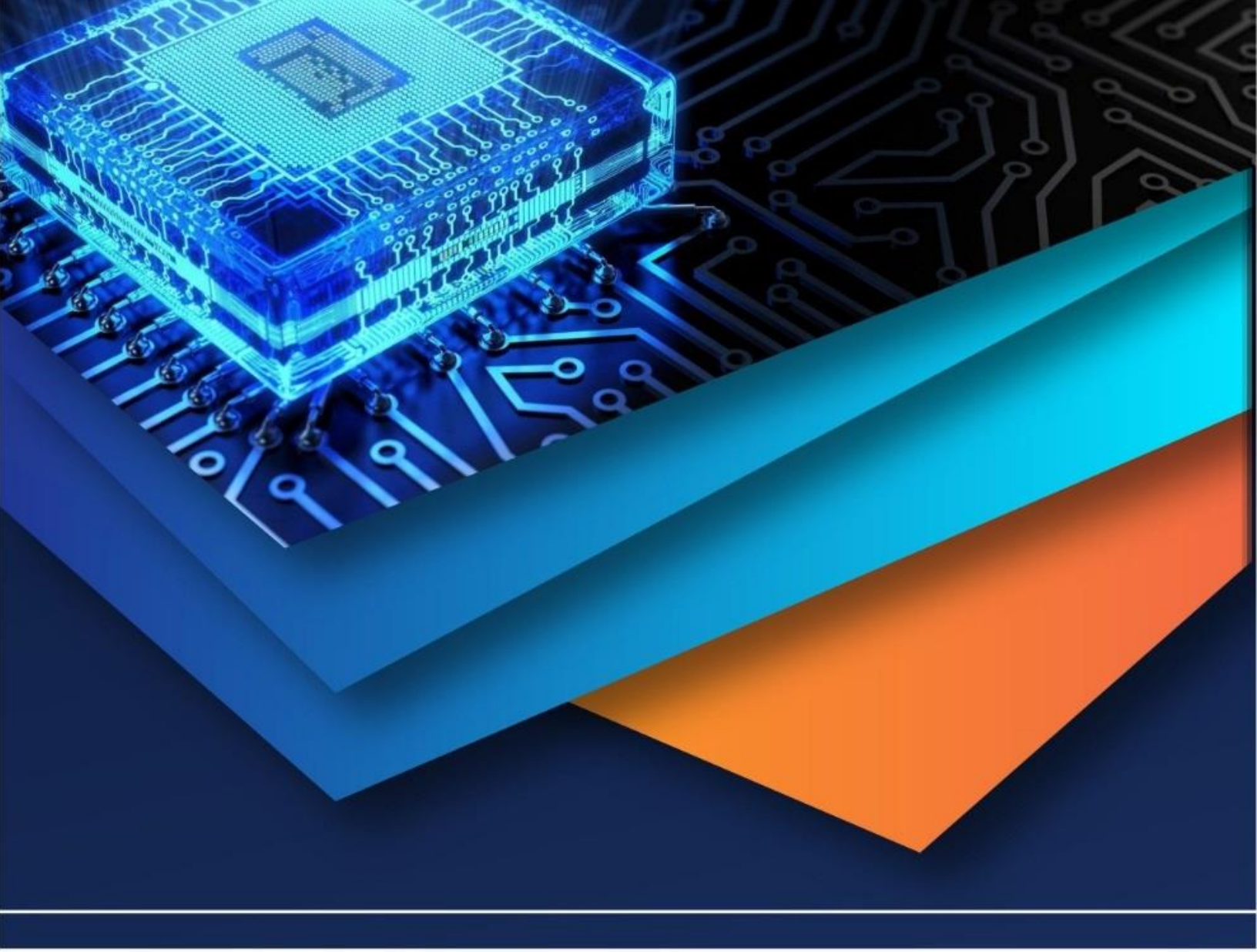
To study and predict the future risk of diabetes disease, the proposed system uses "Multilayer perceptron and Long short term memory." One can keep track of the information that patients provide for future use (to improve the accuracy of the model). The work can also include a visitor enquiry module, in which visitors can submit questions to the administrator, and the administrator can respond. It can also include a treatment section where doctors can upload and patients may check treatment specifics for their cases.

REFERENCES

[1] Research article by Amine Rghioui, Jaime Lloret, Sandra Sendra and Abdelmajid Oumnad :Smart Architecture for Diabetic Patient Monitoring Using Machine Learning Algorithms.
[2] Research Article by:Md. Iftekharul Alam Efat , Shoaib Rahman , Tasnim Rahman:IOT Based Smart Health Monitoring System for Diabetes Patients Using Neural Network.
[3] Research article by Ganjar Alfian , Muhammad Syafrudin , Muhammad Fazal Ijaz , M. Alex Syaekhoni, Norma Latif Fitriyani and Jongtae Rhee :A Personalized Healthcare Monitoring System for Diabetic Patients by Utilizing BLE-Based Sensors and Real-Time Data Processing.
[4] Internet of Things Based on Electronic and Mobile Health Systems for Blood Glucose Continuous Monitoring and Management
[5] Research Article by:Dr Rajendra Kumar Bhartia , Dilip Kumar J Sainib , Somanjoli Mohapatrac , Y.K.Ahmedd , Dr. Sushma Jaiswale and T.M.Nithya:An IoT and fog computing enabled intelligent healthcare monitoring system to optimize the cloud storage.
[6] Research article by Prajoona Valsalan, Tariq Ahmed Barham Baomar, Ali Hussain Omar Baabood :Iot Based Health Monitoring System.
[7] MEDICAL MANAGEMENT OF HYPERGLYCEMIA IN TYPE 2 DIABETES: A CONSENSUS ALGORITHM FOR THE INITIATION AND ADJUSTMENT OF THERAPY
[8] AN IOT BASED DIABETIC PATIENTS MONITORING SYSTEM USING MACHINE LEARNING AND NODE MCU

[9] Research article by Hindawi: Machine Learning Based Diabetes Classification and Prediction for Healthcare Applications.
[10] Research article by Farman Ali, Shaker El-Sappagh, S.M. Riazul Islam, Amjad Ali , Muhammad Attique, Muhammad Imran, Kyung-Sup Kwak: An intelligent healthcare monitoring framework using wearable sensors and social networking data.
[11] Mobile Health Monitoring: Development and Implementation of an app in a Diabetes and Hypertension Clinic
[12] Research article by Ganjar Alfian, Muhammad Syafrudin, Muhammad Anshari, Filip Benes, Fransiskus Tatas Dwi Atmaji, Imam Fahrurrozi, Ahmad Fathan Hidayatullah, Jongtae Rhee: Blood Glucose Prediction Model for Type 1 Diabetes based on Extreme Gradient

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)