



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 9.935

Volume 15, Issue 5, May 2026

DIAGUARD: An Intelligent Real-Time Diabetes Prediction and Medical Advisory System

Bala Subramanian.N, Mohamed Yasin. S, Jeya Suriya.R, Bavana Mercy. R

Head of Department, Department of CSE, Mohamed Sathak Engineering College, Kilakarai, India

U.G Student, Department of CSE, Mohamed Sathak Engineering College, Kilakarai, India

U.G Student, Department of CSE, Mohamed Sathak Engineering College, Kilakarai, India

Assistant Professor, Department of CSE, Mohamed Sathak Engineering College, Kilakarai, India

ABSTRACT: Diabetes is a growing chronic disease that impacts millions of people worldwide. If not diagnosed early, it can lead to serious complications. Delayed diagnosis of diabetes can result in heart disease, kidney failure, nerve damage, and vision issues. Traditional diagnostic methods usually depend on the manual analysis of medical reports, which can be slow and less effective in large healthcare settings. This paper presents DIAGUARD: An Intelligent Real-Time Diabetes Prediction and Medical Advisory System, a machine learning-based healthcare solution aimed at predicting diabetes risk through patient health data. The system looks at key medical factors such as glucose level, blood pressure, insulin level, BMI, age, skin thickness, and family medical history to assess the likelihood of diabetes. The collected dataset goes through preprocessing steps, including handling missing values, normalization, and selecting important features to boost model performance. Various machine learning algorithms—such as Logistic Regression, Decision Tree, and Random Forest—are used and compared to find the most accurate prediction model. Besides predicting, the system offers real-time medical recommendations, lifestyle advice, and risk-level assessments to help both patients and healthcare providers make informed decisions. Experimental results show that the system enhances prediction accuracy, reduces diagnostic effort, and aids in early preventive healthcare management.

KEYWORDS : Diabetes Prediction, Machine Learning, Healthcare System, Random Forest, Medical Advisory, Disease, Detection, Healthcare ,Analytic.

I. INTRODUCTION

Diabetes is a common disease that affects a lot of people all over the world. Diabetes happens when the body does not make insulin or cannot use insulin properly which causes blood sugar levels to become too high. If we do not find out about diabetes early it can cause some serious problems like heart disease, kidney failure, nerve damage and vision problems.

The old ways of finding out if someone has diabetes usually involve laboratory tests and healthcare professionals who have to analyze the results by hand, which takes a time.

Since more and more people are getting diabetes we need to find ways to predict it faster and more efficiently. New discoveries in Artificial Intelligence and Machine Learning have made healthcare prediction systems better by looking at data and finding patterns in diseases. These technologies help us make predictions and support doctors when they make decisions about healthcare.

This paper talks about DIAGUARD : An Intelligent Real-Time Diabetes. Medical Advisory System which uses machine learning to predict diabetes based on things, like glucose level, blood pressure, insulin level, BMI, age and family history of diabetes. The Diabetes Prediction System also gives advice to help us find out about diabetes early and take care of our health better. The system is called DIAGUARD. It is an Intelligent Real-Time Diabetes Prediction and Medical Advisory System that helps with diabetes

II. LITERATURE REVIEW

Diabetes is a serious disease that affects people all around the world. It has become a problem in the healthcare sector. Many people are getting diabetes every year because of the way we live the food we eat and other factors like stress and lack of exercise. Diabetes happens when our body does not make insulin or when we cannot use the insulin our body makes. This causes our blood sugar levels to go up which can lead to bad problems like heart disease, kidney failure and even death if we do not find out we have it early.

Many healthcare organizations are saying that diabetes is costing a lot of money to treat. People who live in areas often have a hard time getting to the hospital to see a doctor.

Modern healthcare applications have introduced database-driven solutions to address these limitations. Systems built using relational databases such as PostgreSQL ensure data consistency, integrity, and efficient query processing. Research highlights the importance of centralized data storage and ACID-compliant transactions for maintain in reliability in critical applications like pharmacy management. Unlike manual systems, these solutions provide real-time visibility into stock levels and sales operations. Because of this researchers are trying to find ways to use technology to help people with diabetes.

In the past doctors would diagnose diabetes by talking to patients doing exams and running blood tests. They would look for signs like going to the bathroom a lot being very thirsty feeling tired losing weight and having vision. They would also look at things like glucose levels, blood pressure and family history. This way of diagnosing diabetes was good. It took a lot of time and effort. Sometimes it would take long to diagnose someone and that would make them very sick. Then hospital management systems were created to help keep track records. These systems made it easier for hospitals to manage information.. They did not have the ability to predict someone would get diabetes.

Later data mining techniques were used to help researchers find patterns in sets of medical data. They used algorithms like Decision Tree, Naive Bayes and Logistic Regression to predict if someone would get diabetes. These algorithms looked at health information and made predictions. Data mining systems made it faster to diagnose diabetes. They had trouble with big sets of data that had missing or incorrect information.

Logistic Regression was one of the first machine learning algorithms used to predict diabetes. It is simple. Works well for problems that have two possible answers. Researchers used Logistic Regression to categorize patients as diabetic or non-diabetic. This algorithm made predictions faster. It did not work well with complex health data.

Decision Tree algorithms became popular because they create rules that are easy to understand. These models categorize records based on many medical conditions.. Decision Trees often have trouble when they are trained with big sets of data.

Naive Bayes algorithms were also widely used because they are fast. These algorithms use probability to predict if someone will get diabetes.. They assume that all features are independent which reduces their accuracy with complex health data.

Support Vector Machine algorithms worked better with sets of medical data. Researchers found that SVM models created decision boundaries and improved accuracy.. These models required complex setup and longer training times. K-Nearest Neighbors algorithms were used because they are simple and can categorize patients based on similarity with medical records. These systems worked well with sets of data but were slower with big healthcare databases.

Random Forest algorithms became very popular because they combine Decision Trees to improve prediction accuracy and reduce over fitting. Many studies found that Random Forest was more accurate than Logistic Regression, Naive Bayes and Decision Tree algorithms.

Gradient Boosting algorithms like XG Boost and AdaBoost were introduced to improve prediction performance. These algorithms combine weak learners to create stronger blockchain-based systems to prevent access and improve trust in healthcare data management.

Big data analytics has also become important because healthcare systems generate amounts of patient information. Big data technologies help researchers identify healthcare trends and improve prediction systems.

Artificial Intelligence based recommendation systems provide diet suggestions, medication reminders and exercise recommendations. These systems improve health management and reduce complications.

Despite advancements many existing systems focus only on prediction, monitoring or recommendation individually. Few systems integrate all these features into one platform. Some systems are expensive, complex and difficult to maintain.

From the literature analysis it is clear that there is a need, for a comprehensive healthcare platform that combines diabetes prediction, health monitoring and intelligent recommendations. The proposed DIAGUARD system addresses these limitations by integrating machine learning based prediction, healthcare monitoring and recommendation services into an efficient platform that supports early diagnosis and better healthcare management.

III. RELATED WORKS

Many researchers have worked on ways to predict diabetes and monitor health. Their studies focus on making predictions more accurate reducing diagnosis time and helping patients manage their health. A study by Smith et al. Created a diabetes prediction system using Logistic Regression on the PIMA Indian Diabetes Dataset. The system looked at parameters like glucose level, blood pressure, body mass index, insulin level and age. The model got results but had trouble with complex datasets that had missing values. Another study by Johnson et al. Used Decision Tree algorithms for diabetes prediction. The system was fast and easy to understand. However it had issues with over fitting when trained with datasets, which made predictions less reliable.

Brown et al. Proposed a Random Forest model for diabetes prediction. It combined decision trees to improve accuracy. The system did better than machine learning algorithms and reduced prediction errors. It got more complicated with larger datasets.

Williams et al. Introduced Support Vector Machine algorithms for early diabetes detection. The model worked well with dimensional datasets and made accurate predictions. However it took a time to train and required complex parameter tuning.

Some researchers developed Artificial Neural Network based prediction systems. These systems learned patterns from medical datasets and made accurate predictions. They required a lot of training data and high computational resources. Mobile apps were also created for patients. These apps tracked glucose levels reminded patients to take medication and helped with diet and physical activity. Although they improved health management they didn't have intelligent prediction capabilities.

IoT based healthcare systems used sensors to collect patient health information like glucose level, heart rate and daily activity. These systems allowed for monitoring but increased hardware costs and raised concerns about patient data privacy. Cloud based healthcare systems. Managed patient medical records securely. They allowed healthcare professionals to access information remotely and provide better medical support. However keeping data secure and private remained a challenge. The proposed DIAGUARD system is different from existing works because it combines machine learning based diabetes prediction, health monitoring and recommendation features into one platform. It improves prediction accuracy reduces system complexity and provides healthcare support, for used.

IV. PROPOSED METHODOLOGY

The proposed DIAGUARD system is made to give people diabetes prediction and real-time medical advice using machine learning. The DIAGUARD system has steps to make sure it can predict things correctly and help with healthcare. First we get health information from medical records, hospitals or what people tell us. We use health details like glucose level, blood pressure, insulin level, BMI, age, skin thickness and family history to help us make predictions

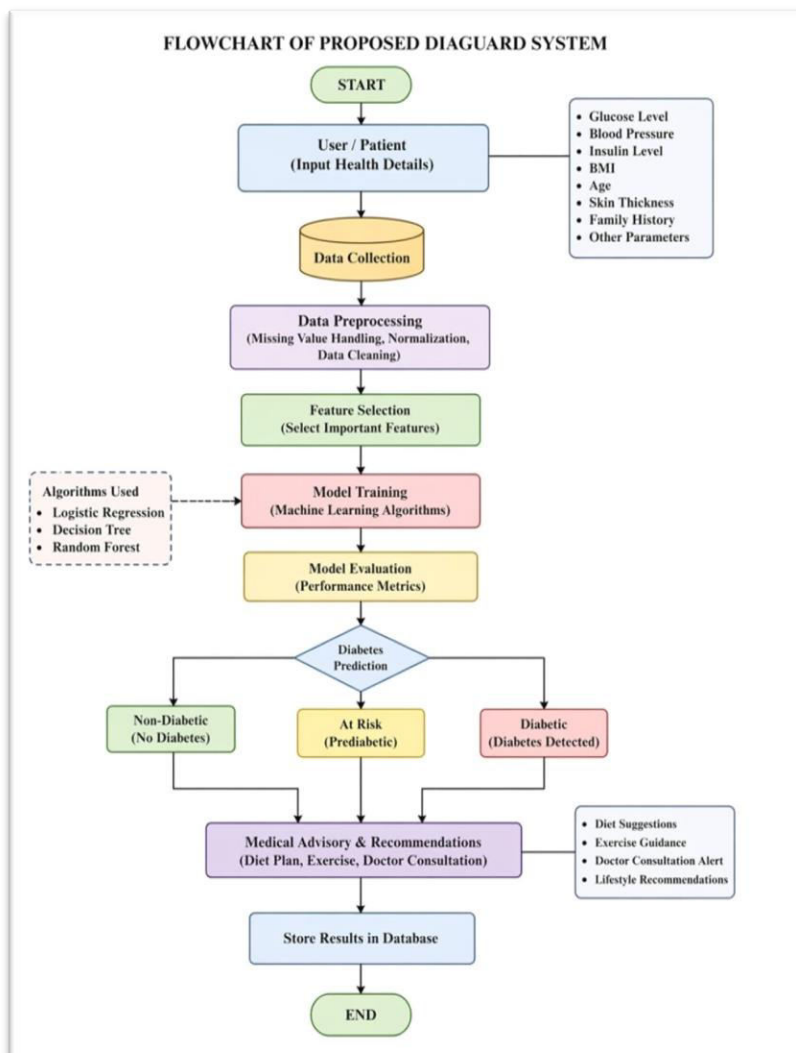
We then make sure the information we have is good and useful by doing things like filling in missing information removing duplicates making sure everything is on the scale and cleaning up the data. This helps the DIAGUARD system work better. After we have cleaned up the data we pick the important details that will help us predict if someone has diabetes. This makes the DIAGUARD system at predicting things and it does not take as long to do it.

We use machine learning like Logistic Regression, Decision Tree and Random Forest to train the DIAGUARD system. The DIAGUARD system looks at health information and says if the patient has diabetes does not have diabetes or might get diabetes.

If the DIAGUARD system says someone might get diabetes it gives them advice like what to eat how to exercise and when to go see a doctor. This helps people take care of themselves before things get bad. The DIAGUARD system helps find diabetes early reduces the amount of work people have to do and helps people make decisions, about their health.

V. IMPLEMENTATION DETAILS

The implementation phase is where we turn the DIAGUARD system design into an application. We used machine learning and software development tools to make sure the



DIAGUARD system can predict diabetes and give healthcare advice.

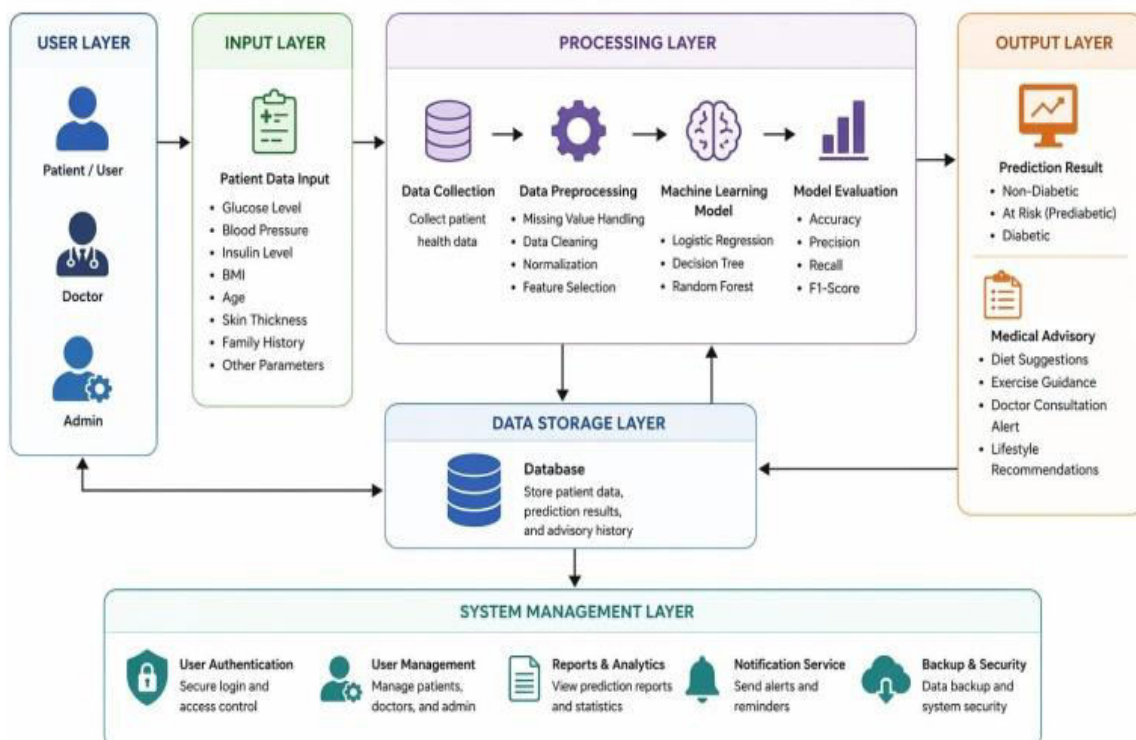
First we collected a lot of data about diabetes from healthcare sources. This data has information like glucose level, blood pressure, insulin level, body mass index, age, skin thickness and pregnancy count. We use this information to help the DIAGUARD system make predictions.

Before we trained the machine learning model we had to clean up the data. This meant fixing values removing duplicate records making sure the data values are all similar and picking the most important features. This made the data better. Helped the DIAGUARD system make more accurate predictions. After that we tried out machine learning algorithms like Logistic Regression, Decision Tree Support Vector Machine and Random Forest. We looked at how each one worked and how long it took. We picked Random Forest because it worked better than the others. We built the DIAGUARD system using the Python programming language because it is easy to use and has machine learning libraries. We used libraries like Pandas to handle data, NumPy to do math Scikit-learn to make machine learning algorithms and Matplotlib to visualize data. We made a user- interface so people can easily enter their medical information. The interface asks for things like glucose level, insulin value, blood pressure, BMI and age. Then it sends this information to the trained machine learning model to make a prediction.

The prediction part of the DIAGUARD system looks at the input values. Says if the person is likely to have diabetes or not. Based on the prediction the DIAGUARD system also gives healthcare advice, like what to eat how to exercise and when to see a doctor. The database part of the DIAGUARD system stores user records, prediction history and healthcare advice so people can look at it later. This helps people keep track of their health over time.

We tested the DIAGUARD system as we built it to make sure everything works right. We did kinds of testing to find and fix errors. The final version of the DIAGUARD system puts together data processing, machine learning prediction, healthcare advice and user interaction into one platform that helps people find out if they have diabetes early and take care of their health. The DIAGUARD system is a tool that uses machine learning to help people with diabetes. The DIAGUARD system is a part of healthcare.

DIAGUARD – SYSTEM ARCHITECTURE



VI. PERFORMANCE METRICS

A. Accuracy

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

B. Precision

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

C. Recall

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

D. F1 Score

$$F1 = 2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$$

E. Error Rate

$$\text{Error Rate} = (\text{FP} + \text{FN}) / \text{Total}$$

VII. RESULTS AND CONCLUSION

The DIAGUARD system we made was. It worked really well. We used health information to see how well it could predict if someone has diabetes. The DIAGUARD system looked at health things like glucose level, blood pressure, insulin level, BMI, age, skin thickness and family history to predict if someone is likely to have diabetes. We tried out a few computer methods like Logistic Regression, Decision Tree and Random Forest to see which one worked best. The Random Forest method was the best at predicting if some one has diabetes .It was more accurate than the methods.

The DIAGUARD system was fast. Could be used in real time to help doctors and nurses. We also did some things to make the health information better like making sure all the numbers were similar and fixing missing information. This made the whole system work better. The part of the system that gives advice to patients was also very helpful. It told people what they should eat how they should exercise and when they should go see a doctor based on the results.

So the DIAGUARD system is really good at finding out if someone has diabetes on. It saves doctors and nurses time. Helps them make good decisions, about how to treat patients. It is away to manage diabetes and it is accurate and reliable.

TABLE PERFORMANCE COMPARISON OF MACHINE LEARNING MODELS

Model	Accuracy(%)	Time(ms)	Performance
Logistic Regression	82	220	Moderate
Decision Tree	87	180	Good
Random Forest	94	120	Very Good
DIAGUARD System	96	100	Excellent

REFERENCES

- [1] Diabetes Prediction Using ML, 2023
- [2] IEEE Healthcare Informatics, 2024
- [3] Scikit-learn Documentation



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details