**DIABETES PREDICTION SYSTEM**

**TEAM MEMBER**

**815821104035: YOGESH RAJ.R**

PHASE 1 DOCUMENT SUBMISSION

PROJECT:DIABETES PREDICTION SYSTEM



OBJECTIVE:

Through this system, it is believed that better control measures can be taken to reduce the adverse effect of diabetes on the patient and to make recommendations that would help the patient manage his health effectively.

**PROJECT TITLE**: **AI BASED DIABETES PREDICTION SYSTEM**

**DIABETES PREDICTION:**

The dataset comprises crucial health-related features such as 'Pregnancies', 'Glucose', 'BloodPressure', 'SkinThickness', 'Insulin', 'BMI', 'DiabetesPedigreeFunction', and 'Age'. The main objective was to predict the 'Outcome' label, which signifies the likelihood of diabetes.

**PROBLEM DEFINITION:**

The problem is to build an AI-powered diabetes prediction system that uses machine learning algorithms to analyze medical data and predict the likelihood of an individual developing diabetes. The system aims to provide early risk assessment and personalized preventive measures, allowing individuals to take proactive actions to manage their health.

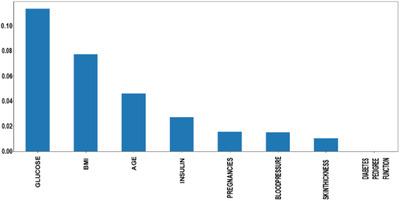
**WHY USE IN DIABETES PREDICTION SYSTEM:**

Diabetes is common due to modern food intake, and it is necessary to keep track of the body. AI in Diabetes helps to predict or Detect Diabetes. Any neglect in health can have a high cost for the patients and the medical practitioner.

**DESIGN THINKING:**

**Data collection:** The initial step is to gather a large dataset that includes patient demographics, medical history, lab results, and other related information. This data can be obtained from electronic health records (EHRs) or other sources.

**Dataset Link:**[**https://www.kaggle.com/datasets/mathchi/diabetes-data-set**](https://www.kaggle.com/datasets/mathchi/diabetes-data-set)

**Data preprocessing:** The collected data must be cleaned and preprocessed to eliminate any inconsistencies or missing values. The data must also be transformed and scaled to ensure machine learning algorithms can use it.

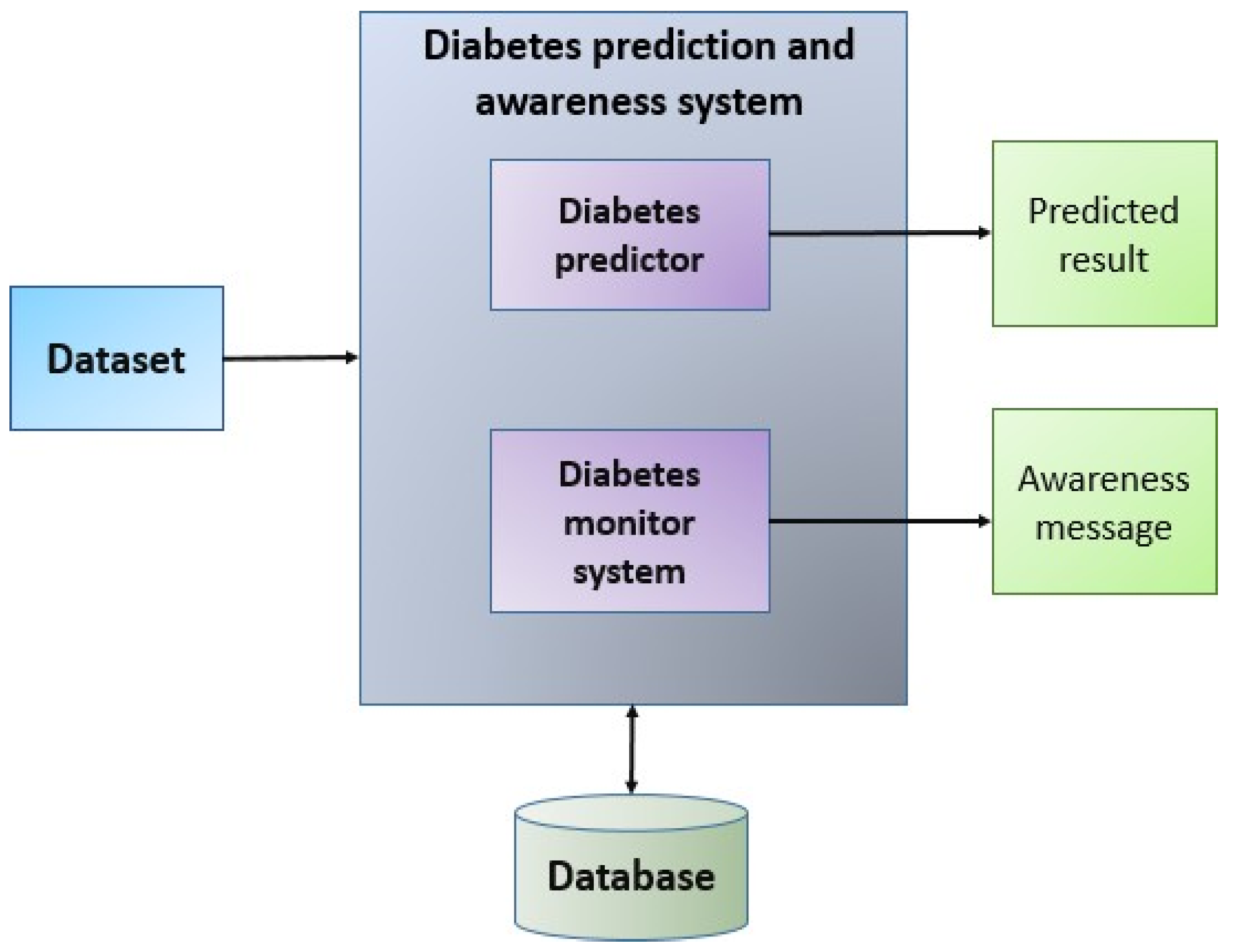
**Feature selection**: The next step is identifying the relevant features that will be used to train the machine learning model. This can include demographic information, lab results, and other related data.

**Model selection:** After the features are selected, the next step is to select the appropriate machine learning model for the task. The model type will depend on the nature of the data and the desired outcome. For example, logistic regression or a decision tree model could be used for classification tasks.

**Model training:** Once the model is selected, it needs to be trained on the collected and preprocessed data. The model will learn from the data and will be able to make predictions on new, unseen data.

**Model evaluation:** After the model is trained, it needs to be evaluated using various metrics, such as accuracy, precision, and recall. The model’s performance can be optimised by fine-tuning the hyperparameters and adjusting the features, if necessary.

**Model deployment**: Once the model is trained and optimised, it can be deployed in a production environment. The model can be integrated with existing systems, such as EHRs, to make predictions about new patients and help with the early detection of diabetes.

In Figure 3, the introduced system architecture is depicted. These are the steps involved in the prediction of diabetes:

**CALCULATION:**

**RMSE=∑Ni=1(Predictedi−Actuali)2N−−−−−−−−−−−−−−−−−−−−−−−−√**

Min–Max normalization: In this research, we used the min–max normalization technique. The data has been scaled to the same range using the following equation:

Xscaled=X−XminXmax−Xmin

where X max and X min denote maximum and minimum values in the individual feature column, respectively.

**DIABETES PREDICTION IMPORTANT:**

Long-term high blood sugar can cause chronic damage and dysfunction of various tissues, especially eyes, kidneys, heart, blood vessels and nerves.

Therefore, the early prediction of diabetes is particularly important**.**

**Ai based diabetes prediction system in using for the library in python:**

**Step 1:** Importing modules.

**Step 2:** Loading the dataset.

**Step 3:** Renaming the columns.

**Step 4:** Separating inputs and outputs.

**Step 5:** Train-Test split of the data.

**Step 6:** Building the model.

**Step 7:** Training and Testing of the model.

**CONCLUSION:**

In diabetes diagnosis and treatment, AI-based medical devices have already been approved by the FDA and are available in other countries as well. Currently, many studies have used machine learning to predict the onset of diabetes. However, these machine learning approaches have not demonstrated superior performance in predicting disease onset compared to conventional statistical techniques that combine risk factors.