

Phase 1: Problem Definition and Design Thinking.

## AI-Based Diabetes Prediction System

Develop an AI-powered diabetes prediction system that leverages machine learning algorithms to analyze medical data and predict the likelihood of an individual developing diabetes, providing early risk assessment and personalized preventive measures.

### Abstract:

Diabetes mellitus is a chronic metabolic disorder affecting millions of individuals worldwide, posing significant healthcare challenges. Early detection and management of diabetes can greatly improve patients' quality of life and reduce healthcare costs. In this context, this paper presents an innovative AI-based chatbot system developed using Python to predict the risk of diabetes in individuals.

The proposed chatbot leverages state-of-the-art machine learning algorithms, including deep neural networks, to analyze a wide range of input data such as personal health records, lifestyle choices, and genetic predispositions. Users interact with the chatbot through natural language conversations, making it accessible to a broad audience.

The chatbot employs a multi-step process to provide predictive insights:

1. **Data Collection:** Users are prompted to input relevant health and lifestyle information, including age, gender, family history, dietary habits, physical activity, and medical history.
2. **Data Preprocessing:** The chatbot preprocesses the input data, performing tasks such as data normalization and feature scaling to ensure model compatibility.
3. **Machine Learning Model:** A trained machine learning model, such as a deep neural network or a decision tree classifier, is used to predict the likelihood of diabetes based on the user's input.
4. **Risk Assessment:** The chatbot interprets the model's predictions and provides the user with a personalized risk assessment, including the probability of developing diabetes in the near future.
5. **Recommendations:** Depending on the risk level, the chatbot offers tailored recommendations for lifestyle modifications, regular check-ups, and potential medical interventions.

This AI-based chatbot is designed to empower individuals with the knowledge and guidance they need to take proactive measures in managing their health. The system's user-friendly interface, natural language processing capabilities, and accurate predictive modeling contribute to its potential as a valuable tool for diabetes prevention and management.

## PROPOSED METHODOLOGY

Goal of the paper is to investigate for model to predict diabetes with better accuracy. We experimented with different classification and ensemble algorithms to predict diabetes. In the following, we briefly discuss the phase.

### Dataset Description-

the data is gathered from UCI repository which is named as Pima Indian Diabetes Dataset. The dataset have many attributes of 768 patients.

### Dataset Description

Attributes

- 1 Pregnancy
- 2 Glucose
- 3 Blood Pressure
- 4 Skin thickness
- 5 Insulin
- 6 BMI(Body Mass Index)
- 7 Diabetes Pedigree Function
- 8 Age

The 9th attribute is class variable of each data points. This class variable shows the outcome 0 and 1 for diabetics which indicates positive or negative for diabetics.

### Distribution of Diabetic patient-

We made a model to predict diabetes however the dataset was slightly imbalanced having around 500 classes labeled as 0 means negative means no diabetes and 268 labeled as 1 means positive means diabetic.

### Algorithm-

- Select the hyper plane which divides the class better.
- To find the better hyper plane you have to calculate

the distance between the planes and the data which is called Margin.

- If the distance between the classes is low then the

chance of miss conception is high and vice versa.  
So we need to

- Select the class which has the high margin.  
Margin = distance to positive point + Distance to negative point.

