**DIABETES RISK PREDICTION**

**1) Installation Procedure:**

* Ensure Python 3.x is installed.
* Install the required libraries using the following commands:

**Execute these commands in terminal:**

pip install streamlit

pip install mysql-connector-python

pip install pandas numpy scikit-learn matplotlib seaborn

* Verify that MySQL is installed and accessible.

**2) Dataset and its Details:**

The dataset used for this project is the Pima Indians Diabetes Dataset, which contains the following features:

* Pregnancies: The number of times the individual has been pregnant. This feature is crucial as gestational diabetes can increase the risk of developing diabetes later in life.
* Glucose: The plasma glucose concentration measured during a 2-hour oral glucose tolerance test. High glucose levels are a key indicator of diabetes.
* Blood Pressure: The diastolic blood pressure (in mm Hg) recorded when the individual was at rest. Abnormal blood pressure levels can be associated with metabolic disorders, including diabetes.
* Skin Thickness: The thickness of the triceps skin fold (in mm) measured with calipers. This measurement helps estimate body fat and insulin resistance.
* Insulin: The 2-hour serum insulin (in mu U/ml) levels measured during the glucose tolerance test. Elevated insulin levels can indicate insulin resistance, a precursor to diabetes.
* BMI: The body mass index calculated from the individual's weight and height. A higher BMI is linked to increased diabetes risk.
* Diabetes Pedigree Function: A function that scores the likelihood of diabetes based on family history and genetic factors. This feature provides insight into the genetic predisposition to diabetes.
* Age: The age of the individual (in years). Aging is a significant risk factor for the development of diabetes.
* Outcome: A binary variable indicating whether the individual has diabetes (1) or not (0). This is the target variable for the prediction model.

By integrating these features, the diabetes risk prediction system aims to provide accurate assessments, empowering users to take proactive measures in managing their health.

**3) Implementation and Results:**

* **Implementation**:
  + The project uses Streamlit to build an interactive web application for predicting diabetes risk.
  + MySQL is used for storing and retrieving data.
  + The predictive model utilizes libraries like Pandas, NumPy, and Sklearn for data analysis and prediction.
  + Fitness-related data is analyzed to derive insights related to the diabetes risk.
* **Results**:
  + The model outputs the risk of diabetes based on inputs related to physical activities and fitness levels.
  + Visualizations are provided using Matplotlib and Seaborn to present trends and predictions.

















