**Diabetes Progression Prediction Using Linear Regression**

**1. Introduction**

Diabetes is a chronic condition that requires continuous monitoring and prediction of disease progression. This project uses the **Diabetes dataset** from sklearn.datasets to develop a **Linear Regression** model that predicts the progression of diabetes based on various patient health indicators.

**2. Dataset and Preprocessing**

The **Diabetes dataset** consists of **442 samples** with **10 numerical features**, including:

* **age**: Age of the patient
* **sex**: Gender of the patient
* **bmi**: Body mass index
* **bp**: Blood pressure
* **s1-s6**: Blood serum measurements

The target variable, **diabetes\_progression**, represents a **quantitative measure of disease progression one year after baseline**.

The dataset was split into **80% training data** and **20% test data** using train\_test\_split().

**3. Model Training**

We implemented **Linear Regression** using LinearRegression() and trained it on the **training dataset**.

**4. Model Evaluation**

The model was tested on the unseen **test dataset**, and its performance was measured using the **R² score**, which indicates how well the model explains the variance in diabetes progression.

* **R² Score:** The model achieved a maximum accuracy of **57%**, meaning it explains **57% of the variance** in diabetes progression.