# Alnd Srud Mirza --- 01-22-00127

# Diabetes Dataset Report

This document provides an overview of the dataset used in our diabetes prediction application. The dataset includes clinical and demographic features collected from 1000 patients, which are used to train a machine learning model for classifying the likelihood of diabetes.

## Dataset Description

The dataset contains 14 columns, including patient demographics and medical test results. The target variable is 'CLASS', which indicates the diabetes category of a patient:  
- N: Non-Diabetic  
- P: Pre-Diabetic  
- D: Diabetic

Features include:  
- ID and No\_Pation: Unique identifiers for each patient  
- Gender: 'M' for Male, 'F' for Female  
- AGE: Age of the patient  
- Urea, Cr, HbA1c, Chol, TG, HDL, LDL, VLDL, BMI: Results from various blood and biochemical tests

## Dataset Statistics

Number of Rows: 1000

Number of Columns: 14

## Example Records

Here are some example patient records from the dataset:

Example 1: AGE: 50, Gender: F, HbA1c: 4.9, BMI: 24.0, CLASS: N

Example 2: AGE: 26, Gender: M, HbA1c: 4.9, BMI: 23.0, CLASS: N

Example 3: AGE: 50, Gender: F, HbA1c: 4.9, BMI: 24.0, CLASS: N

## Usage

This dataset was used to train a machine learning model that powers the GUI application. The model predicts the risk of diabetes based on user input.