

Diabetes Health Indicators

Diabetes mellitus is one of the most serious chronic *illnesses* in the world. Diabetes is a serious chronic disease in which individuals lose the ability to effectively regulate levels of glucose in the blood, and can lead to reduced quality of life and life expectancy.

The goal of the project is to:

- Predict if the user has prediabetes/diabetes or no diabetes from the given features.

Main steps:

Preprocessing:

- Data Collection.
- Data Cleaning/ Cleansing.

Feature Selection and Extraction:

- Using any feature selection methods.
- (optional) Visualize correlation.

Model Training:

- Train using **Logistic Regression**, **SVM** and **Decision Tree** models.

Model Evaluation:

- Accuracy
- Precision
- F1 score
- Confusion matrix

Test the models with new data

Dataset Description:

- diabetes _ binary _ health _ indicators _ BRFSS2015.csv
dataset of 253,680 survey responses to the CDC's BRFSS2015.

- The target variable Diabetes_binary has 2 classes. 0 is for no diabetes, and 1 is for prediabetes or diabetes.

This dataset has 21 feature variables and is not balanced.

Dataset: Diabetes Health Indicators