Project Title: Diabetes Prediction using Machine Learning on PIMA Indian Diabetes Dataset provided by UCI Machine Learning Datasets.

Author:

Date:

Objective:

We will try to build a machine learning model to accurately predict whether or not the patients in the dataset have diabetes or not based on learning the features and patterns from the dataset provided, thus evaluating the performances of given Machine Learning Models based on unseen dataset.

Context:

Diabetes, is a group of metabolic disorders in which there are high blood sugar levels over a prolonged period. Symptoms of high blood sugar include frequent urination, increased thirst, and increased hunger. If left untreated, diabetes can cause many complications. Acute complications can include diabetic ketoacidosis, hyperosmolar hyperglycemic state, or death. Serious long-term complications include cardiovascular disease, stroke, chronic kidney disease, foot ulcers, and damage to the eyes

About the Dataset:

This dataset is originally from the National Institute of Diabetes and Digestive and Kidney Diseases. Several constraints were placed on the selection of these instances from a larger database. In particular, all patients here are females at least 21 years old of Pima Indian heritage.

Technical Solution Overview:

With extensive application of Supervised Machine Learning algorithms on the Data set, our aim is to reach the maximum accuracy on unseen dataset by comparative study of Naïve Bayes , Logistic Regression performances.

Similar Solution:

There are many viable methods to prevent and the warn the patients about their medical status based on reports and experts observations, However Machine Learning Model makes it easier to identify at the early stage and also with less factor being involved.

Project Work Breakdown:

A Machine Learning project cycle follows the below steps:

1. Exploratory Data Analysis

2.Data Cleaning

3.Feature Engineering

4.Model Building

5.Model Evaluation

6.Conclusion