GLA University, Mathura-2021

Diabetes Prediction

(Mini project Synopsis)

Under the supervision

Of

Mr. Amir Khan
(Technical Trainer)
Department of Computer Science and Engineering
GLA University, Mathura

Efforts By:
Janvi Pangoriya (181500292) Section D
Nidhi Gupta (181500422) Section D

1.1 GROUP DISCUSSION

- **Janvi Pangoriya**
  University Roll No: 181500292
  Contact No.: 7617740300
  Mail Id: janvi.pangoriya_cs18@gla.ac.in
  Github Id: https://github.com/JanviPangoriya

- **Nidhi Gupta**
  University Roll No:181500422
  Contact No: 6396492523
  Mail Id: nidhi.gupta_cs18@gla.ac.in
  Github Id: https://github.com/nidhi2611

1.2 INTRODUCTION

Proposed Diabetes Prediction is a machine learning application which predicts whether the patient whose details has been entered is diabetic or not. There are various parameters on which this model will predict the condition. Some of them include the number of pregnancies the patient had, the BMI (Body Mass Index), the glucose intake, diastolic blood pressure, skin thickness, age etc.

1.3 MOTIVATION

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.

So it is really important to predict diabetes before any of the disorders take place and this application can be easily implemented through various machine learning classification algorithms that have been discovered and this can be an further asset to medical sciences.

1.4 DATASET

This **dataset** is originally from the National Institute of Diabetes and Digestive and Kidney Diseases. The objective of the dataset is to diagnostically predict whether or not a patient has diabetes, based on certain diagnostic measurements included in the dataset. Several constraints were placed on the selection of these instances from a larger database. The datasets consists of several medical predictor variables and one target variable, Outcome. Predictor variables includes the number of pregnancies the patient has had, their BMI, insulin level, age, and so on.

- **Pregnancies**: Number of times pregnant

- **Glucose**: Plasma glucose concentration a 2 hours in an oral glucose tolerance test
- **Blood Pressure**: Diastolic blood pressure (mm Hg)
- **Skin Thickness**: Triceps skin fold thickness (mm)
- **Insulin**: 2-Hour serum insulin (mu U/ml)
- **BMI**: Body mass index (weight in kg/(height in m)^2)
- **Diabetes Pedigree Function**: Diabetes pedigree function
- **Age**: Age (years)
- **Outcome**: Class variable (0 or 1)

1.5 EXPECTED OUTCOME

After the building of this machine learning model, we will be able to detect whether the person is suffering from diabetes or not as it is necessary for a person to maintain proper health.

1.6 FUTURE SCOPE

Since this dataset is originally from the National Institute of Diabetes and Digestive and Kidney Diseases, this may be taken as a real time application to predict diabetes in a patient suffering from any of the symptoms and can be a contribution to the medical sciences.

A particular method to identify diabetes is not very sophisticated way for initial diabetes detection and it is not fully accurate for predicting diseases. Thats why we need a smart hybrid predictive analytics diabetes diagnostic system that can effectively work with accuracy and efficiency. We can use data mining, neural network for exploring and utilizing to support medical decision, which improves in diagnosing the risk for pregnant diabetes. Due to the dataset we have till date are not up to the mark, we cannot predict the type of diabetes, so in future we aim to predicting type of diabetes and explore it, which may improve the accuracy of predicting diabetes.