

# **DIABETICS PREDICTION SYSTEM BASED ON LIFE STYLES**

## **1 INTRODUCTION**

### **1.1 Overview**

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. In particular, all patients here are females at least 21 years old of Pima Indian heritage.

### **1.2 Purpose**

In this, we need 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. In particular, all patients here are females at least 21 years old of Pima Indian heritage.

The datasets consist of several medical predictor variables and one target variable, Diabetes. Predictor variables include the number of pregnancies the patient has had, their BMI, insulin level, age, and so on.

## **2 LITERATURE SURVEY**

### **2.1 Existing problem**

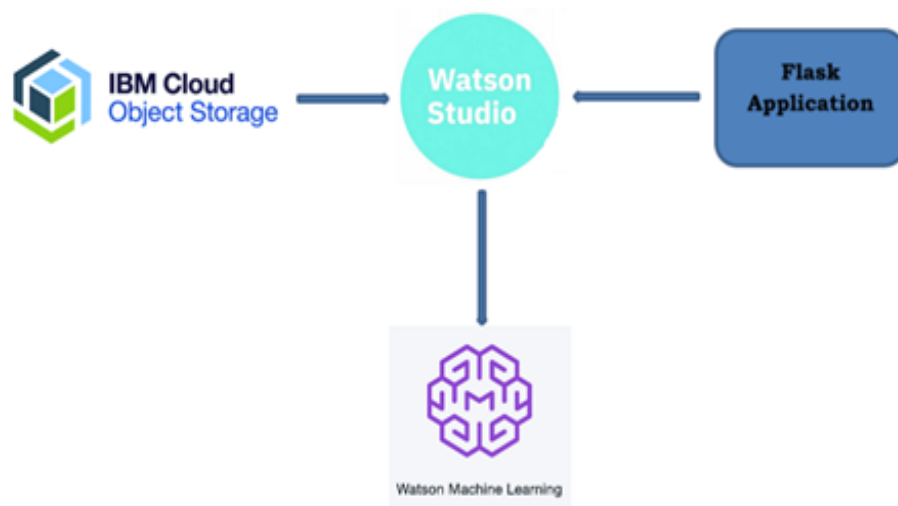
**At present it is very difficult to predict the diabetes accurately.**

### **2.2 Proposed solution**

**Develop an end-to-end web application that predicts the probability of females having diabetes. The application must be built with Python-Flask or Django framework with the machine learning model trained & deployed on IBM Watson Studio**

### 3 THEORITICAL ANALYSIS

#### 3.1 Block diagram



#### 3.2 Hardware / Software designing

##### Hardware

PC / Laptop with atleast 64 bit , 2 GHz , 8 Gb RAM , Windows-10 or higher.

##### Software

IBM Watson Studio, IBM Machine Learning, IBM Cloud Object Storage , IBM Nodered,  
Flask Integration

### 4 EXPERIMENTAL INVESTIGATIONS

Kindly refer to the screen shots of file named- **EXPERIMENTAL INVESTIGATIONS** uploaded on Github.

## 5 FLOWCHART

- Diabetics Prediction System Based On Life Style
  - Dataset Collection
    - Collecting Dataset
  - Create IBM Academic Initiative Account
    - Create IBM Academic Initiative Account
    - Login To IBM Cloud
    - Create Cloud Object Storage Service
    - Download Watson Studio Desktop
    - Create Watson Studio Platform
    - Create Machine Learning Service
  - Train A Model In Watson Studio
    - Create A Project In Watson Studio
    - Upload The Dataset
    - Create Notebook Instance
    - Importing Libraries
    - Impoting Dataset From IBM COS
    - Data Visualization
    - Taking Care Of Missing Data
    - One Hot Encoding
    - Feature Scaling
    - Splitting Data Into Train And Test

- Train The Model Using Classification Algorithms
- Deploying The Model
  - Deploy The Model In Watson Machine Learning
  - Generate Scoring End Point
- Build Flask Application
  - Create An HTML File
  - Build Python Code

## 6 RESULT

Kindly refer to the screen shots of file named- **OUT PUT SCREENSHOTS , Result-0 & Result-1 files** uploaded on Github.

## 7 ADVANTAGES & DISADVANTAGES

Accurate prediction

Faster method-Low time consuming

Precise

Mass data output

## 8 APPLICATIONS

**Diabetes predictions of the patient in different Hospitals.**

## 9 CONCLUSION

**We can build a machine learning model to accurately predict whether or not the patients in the dataset have diabetes or not?**

## 10 FUTURE SCOPE

**We can collect the data of Humans Area wise of whole the city & predict the Diabetes.**

## **11 BIBILOGRAPHY**

### **APPENDIX**

- A. Source code-5-Snippet codes-ZIP
- B. Data file-Diabetes
- C. Jason flow-file
- D. Experimenal Investigation -ZIP
- E. OUTPUT-Screen shots
- F. Output files
- G. Online Project Doc
- H. Video Demonstration link-README.md file

○