BioInformatics Report

Arkaprabha Majumdar
2018MSDBA015
Central University of Rajasthan, India

Report Info:

5 November 2019

pravaa8@gmail.com

Keywords: diabetes, hypoglycema, machine learning, fuzzy computing

Abstract

Diabetes is a serious disease that you cannot treat on your own. Treatment for diabetes requires keeping close watch over your blood sugar levels (and keeping them at a goal set by your doctor) with a combination of medications, exercise, and diet. If you have type 1 diabetes, your pancreas no longer makes the insulin your body needs to use blood sugar for energy. *Microvascular* complications of diabetes are those long-term complications that affect small blood vessels. These typically include retinopathy, nephropathy, and neuropathy. *Macrovascular* complications of diabetes are primarily diseases of the coronary arteries, peripheral arteries, and cerebrovasculature. Early macrovascular disease is associated with atherosclerotic plaque in the vasculature supplying blood to the heart, brain, limbs, and other organs. Late stages of macrovascular disease involve complete obstruction of these vessels, which can increase the risks of myocardial infarction (tissue death), stroke, claudication, and gangrene. Cardiovascular disease (CVD) is the major cause of morbidity and mortality in patients with diabetes. To address all these issues, it is imperative to find a fast automatic diabetes detection system to correctly identify diabetes (both type-I and type-II) for early precautions.

Introduction

Abnormal blood glucose is a modifiable risk factor for CVD and a diagnosis of diabetes substantially increases a person's absolute CVD risk score.
Tests to detect diabetes :

1. Fasting blood glucose
2. Glycated haemoglobin (HbA1c)
3. Oral glucose tolerance test

The signs and symptoms of Type 1 diabetes usually develop quickly, especially in children, over a period of weeks. In babies and young children, the first indication of Type 1 diabetes may be a yeast infection that causes a severe rash. In young children and infants, lethargy, dehydration and abdominal pain also may indicate Type 1 diabetes.

Once the symptoms appear, a blood test generally will reveal very high blood glucose.
Type 2 diabetes can be detected easily during a routine screening exam and blood test. However, it frequently can go undiagnosed for years unless a physician draws a blood sample to check the blood glucose.
In the early stages of Type 2 diabetes, you experience few to no noticeable signs of the disease. As time goes by and the untreated blood glucose continues to rise, symptoms begin.

Prediabetes:

Prediabetes is a health condition that means your blood sugar level is higher than normal, but not yet high enough for you to be diagnosed with diabetes.

Prediabetes usually has no symptoms, but it almost always shows up before you get diagnosed with diabetes.

About 86 million people in the U.S. over age 20 have prediabetes. And doctors see the need to diagnose it more often. Treating it can prevent more serious health problems later on. These range from type 2 diabetes to problems with your heart, blood vessels, eyes, and kidneys.

By the time you're diagnosed with diabetes, many of these problems have already taken hold. Early diagnosis is important. In the early years of pre-diabetes or diabetes, the beta cells are progressively damaged by high blood sugars.Usually by the time diabetes is diagnosed, half of the beta cells are nonfunctional. This can not be reversed so that the beta cells can go back to insulin production. When an early diagnosis of pre-diabetes is made, almost 100 percent of beta cells are functional. If lifestyle changes are made and some diabetes medications are used right away, many beta cells will stay healthy and make blood sugar control easier.

People who have a higher risk of developing pre-diabetes or Type-II diabetes are:

- those overweight, especially in the abdominal area.
- those with a family history of diabetes
- women who have had gestational diabetes
- people with steroid induced hyperglycemia
- or those having hypertension or an abnormal lipid profile

Below are risks associated with Type 2 diabetes:

- Increased risk of heart attack and stroke
- Coronary artery disease, hypertension, dyslipidemia
- Impotence, blindness, amputations and kidney failure

Methods & Techniques

- Machine Learning:
  - Python (Google Colab Notebook)

    - scikit-learn
    - pandas
- Fuzzy Computing:
    - created membership functions
    - fuzzification of crisp diabetes matrix

Data & Observations

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.
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
- **BloodPressure** - Diastolic blood pressure (mm Hg)
- **SkinThickness** - Triceps skin fold thickness (mm)
- **Insulin** - 2-Hour serum insulin (mu U/ml)
- **BMI** - Body mass index (weight in kg/(height in m)^2)
- **DiabetesPedigreeFunction**
- **Age** (years)
- **Outcome** - Class variable (0 or 1) 268 of 768 are 1, the others are 0

Conclusion and Discussion

We have proposed a novel Artificial Bee Colony (ABC) algorithm in which a mutation operator is added to an Artificial Bee Colony for improving its performance. When the current best solution cannot be updated, a blended crossover operator (BLX-α) of genetic algorithm is applied, in order to enhance the diversity of ABC, without compromising with the solution quality. This modified version of ABC is used as a new tool to create and optimize automatically the membership functions and rules base directly from data.
The performances of the proposed method are evaluated through classification rate, sensitivity and specificity values using 10-fold cross-validation method. The obtained classification rate of our method is 84.21% and it is very promising when compared with the previous research in the literature for the same problem.