

Diabetes Prediction System

Project Proposal for COMP 8157-1 – Advanced Database Topics

Team members:

Ashly Chinnu Varghese #110030728

Tincy Thomas #110026212

Manik Mahajan #110026276

Wen Dong #110057395

Brief description:

Diabetes has grown into one of the most popular and devastating chronic diseases in modern society, a growing number of people contract with diabetes due to a variety of factors, like age, diet, weight, lack of sports etc.

This comes out of the idea to build a predictive and preventive system for potential diabetes group, by applying data mining and analysis to comprehensive diabetes datasets of various sources, the system will be able to foresee risks for given personal information before the chronic exacerbates. It uses distributed MongoDB as the unified data center for gathered datasets, which can facilitate and speed up the system development, and opens possibility accommodating other diseases.

The motivation:

First, it is valuable in such a world where diabetes has become one of the most troublesome and parlous threats to human being, accessibility to a quick and easy diagnosing method in an early phase important.

Initial idea to the solution:

Build a machine learning model based on dataset files instead of advanced database, which was also feasible, but it would lack flexibility in terms of extensibility, and would make the development process long and tough because feeding tremendous data to model from reading files takes time.