Project Topic: Diabetes Prediction

Project Category: Life Sciences

Name: Xinxie Wu

SUNet ID: xinxiewu

Department of Computer Science

Stanford University

[xinxiewu@stanford.edu](mailto:xinxiewu@stanford.edu)

Motivation:

Method:

Intended Experiments:

Dataset

1. Pima Indians Diabetes Database: This dataset is widely used for diabetes prediction and is available from the UCI Machine Learning Repository. It contains data on 768 female Pima Indians, including features such as age, BMI, blood pressure, and glucose levels.
2. Diabetes 130-US hospitals for years 1999-2008 Data Set: This dataset contains data on over 100,000 diabetic patients from 130 hospitals in the United States. It includes features such as patient demographics, lab test results, and hospital admission details.
3. Diabetes Data from the National Institute of Diabetes and Digestive and Kidney Diseases: This dataset contains data on 443 patients with diabetes, including features such as age, sex, BMI, and blood pressure.
4. Diabetes Data Set from Kaggle: This dataset is available on Kaggle and includes data on 101,766 diabetic patients. It includes features such as patient demographics, lab test results, and medication details.
5. Indian Diabetes Dataset: This dataset contains data on 768 patients from India, including features such as age, BMI, blood pressure, and glucose levels.

Reference: