DIABETES PREDICTION PROJECT PROPOSAL

Predicting whether a person has diabetes or not.

1 Context

From Frankfurt hospital, Germany we have information of person's health related metrics such as BMI, glucose levels, insulin levels etc with having known the person is having diabetes or not. Having this data can help in making a model for predicting a person is having diabetes or not based on the person's health related metrices.

2 Criteria for success

Data exploration and using different machine learning models for predicting diabetes for getting good accuracy.

3 Scope of solution space

The data file consists of various types of person's health information for making predictions and drawing conclusions.

4 Constraints within solution space

Data set has only 2000 rows of data.
Getting higher accuracy on predicting diabetes is necessary.

5 Stakeholders to provide key insight

Mentor - Chirasmita Mallick
Springboard team

6 Key data sources

CSV file obtained from kaggle website– Contains the information around 2000 observations having 9 columns.

Steps involved in problem solving

- Data Wrangling - Data Collection, Data Organization, Data Definition & Data Cleaning
- Exploratory Data Analysis
- Pre-processing & Training Data Development
- Modeling - Model Evaluation and Optimization