**Predicting Diabetes Mellitus with Machine Learning**

Diabetes mellitus is a chronic disease characterized by hyperglycemia. By 2040, the world’s diabetic patients will reach 642 million, which means that one of the ten adults in the future is suffering from diabetes.  People having diabetes have high risk of diseases like heart disease, kidney disease, stroke, eye problem, nerve damage, etc. The risk factor and severity of diabetes can be reduced significantly if the precise early prediction is possible.

The main Objective is to predict whether the patient has diabetes or not based on various features like Glucose level, Insulin, Age, BMI.Preprocessing consist of  outlier rejection, filling missing values, data standardization, feature selection, and K-fold cross-validation.  For feature selection PCA and minimum redundancy maximum relevance (mRMR) can be used to reduce the dimensionality. We use decision tree, random forest and neural network to predict diabetes mellitus. Five-fold cross validation is used to examine this model. Further it can be implemented as a real time application.

DATASET : pima Indians dataset