# Title:- RB-Bayes algorithm for the prediction of diabetic in Pima Indian dataset

**Abstract:**

Diabetes is a major concern all over the world. It is increasing at a fast pace. People can avoid diabetes at an early stage without any test. The goal of this paper is to predict the probability of whether the person has a risk of diabetes or not at an early stage. This would lead to having a great impact on their quality of human life. The datasets are Pima Indians diabetes and Cleveland coronary illness and consist of 768 records. Though there are a number of solutions available for information extraction from a huge datasets and to predict the possibility of having diabetes, but the accuracy of their mining process is far from accurate. For achieving highest accuracy, the issue of zero probability which is generally faced by naïve bayes analysis needs to be addressed suitably. The proposed framework RB-Bayes aims to extract the required information with high accuracy that could survive the problem of zero probability and also configure accuracy with other methods like Support Vector Machine, Naive Bayes, and K Nearest Neighbor. We calculated mean to handle missing data and calculated probability for yes (positive) and no (negative). The highest value between yes and no decide the value for the tuple. It is mostly used in text classification. The outcomes on Pima Indian diabetes dataset demonstrate that the proposed methodology enhances the precision as a contrast with other regulated procedures. The accuracy of the proposed methodology large dataset is 72.9%.

#### Keywords: Accuracy; Bayes Method; Cross-validation; PIMA Indian data-set; SVM