paper-title

author-namelist

March 3, 2025

Abstract

This study uses machine learning to diagnose diabetes patients. Naive Bayes Classifier, K-Nearest Neighbours, Logistic Regression, Random Forest and XGBoost models were the algorithms used, with XGBoost showing the most promise. The XGBoost model was thus selected and a visual representation of the importance of each attribute was derived. Present models suffer from being unable to differentiate between different types of diabetes due to lack of data points in a single dataset. This model hopes to alleviate this weakness by compounding two datasets, such that one predicts the presence of diabetes, and the other, the type of diabetes.

1 Introduction

2 Literature Review

2.1 Subsection if needed

3 Methodology