Abstract

Diabetes is a serious malady where one has abnormally high blood sugar levels. Despite being so deadly, it is quite common as anyone is susceptible to it. If untreated, it can damage a person’s kidneys, eyes, nerves, and other organs. Genes, environment, and preexisting medical conditions can all affect a person’s odds of developing diabetes. The bottom line is that it can be extremely deadly if discovered late. Thus, it is imperative that researchers devise an accurate diabetes predictor in order to enable early treatment of diabetic people. This paper demonstrates the prediction of diabetes using the Random Forest algorithm on the PIMA Indians Diabetes dataset. Using important data points and features of several healthy and diabetic PIMA Indians, the model predicts the onset of diabetes. The performance of this algorithm is evaluated using metrics like Accuracy, Precision, and Recall. Furthermore, several suggestions to improve the effectiveness of this model are discussed.


Index Terms

Computer Science          Artificial Intelligence

Keywords

Machine learning, Diabetes prediction, Random Forest, PIMA Indians diabetes dataset