MAJOR PROJECT

Akash perla

Project: This dataset is originally from the National Institute of Diabetes and Digestive and Kidney Diseases. The objective of the dataset is to diagnostically predict whether or not a patient has diabetes, based on certain diagnostic measurements included in the dataset. Several constraints were placed on the selection of these instances from a larger database. In particular, all patients here are females at least 21 years old of Pima Indian heritage.

Tools Used: Google Colaboratory

Algorithm: Logistic Regression Algorithm
K Nearest Neighbor Algorithm
SVM Algorithm

Logistic Regression Algorithm: Logistic regression is one of the most popular Machine Learning algorithms, which comes under the Supervised Learning technique. It is used for predicting the categorical dependent variable using a given set of independent variables.

K Nearest Neighbor Algorithm: K-Nearest Neighbour is one of the simplest Machine Learning algorithms based on the Supervised Learning technique. K-NN algorithm can be used for Regression as well as for Classification but mostly it is used for Classification problems.

SVM Algorithm: Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems. However, primarily, it is used for Classification problems in Machine Learning.

The accuracy of the diabetes dataset using the Logistic Regression Algorithm is 0.7532467532467533

The accuracy of the diabetes dataset using the K Nearest Neighbor Algorithm is 0.7532467532467533

The accuracy of the diabetes dataset using the SVM Algorithm is 0.7316017316017316

The best algorithm for the diabetes dataset is Logistic regression Algorithm and K Nearest Neighbor Algorithm because it has more accuracy than theSVM algorithm.