

Date	19 September 2022
Team ID	PNT2022TMID54403
Project Name	Visualizing And Predicting Heart Diaseases With An Interactive Dash Board
Maximum Marks	4 Marks

## **IDEATION**

1. The so-called big data produced by the healthcare sector contains vast amounts of information that can be used to make decisions. In order to create decisions that are more accurate than intuition, a vast amount of data is used. EDA identifies errors, locates pertinent data, verifies presumptions, and establishes the relationship between the explanatory factors. In this context, EDA is understood to be data analysis without statistical modeling or inferences. Any profession needs analytics since it can predict the future and reveal hidden patterns. In the recent past, data analytics has been regarded as a cost-effective technology and it now plays a crucial role in healthcare, including new study discoveries, emergency circumstances, and disease outbreaks. EDA is a crucial step when analyzing data, and the application of analytics in healthcare improves treatment by simplifying preventive care. In this study, the K-means algorithm is used to analyze publically available data on heart disease and to forecast the risk variables that lead to heart disease. The dataset contains 209 records with eight parameters, including age, the type of chest pain, blood pressure, blood sugar level, resting ECG, heart rate, and four different types of chest pain. K-means clustering method, together with data analytics and visualization tools, are utilised to forecast cardiac disease.

2. Early heart disease diagnosis may help somewhat to lower the death rate. This program aids in the early diagnosis and prediction of heart disease. Healthcare businesses now generate massive amounts of data, yet those data are incredibly fragmented. This data can be used to predict cardiac illnesses with simplicity if it is properly organized using data mining techniques. To create a decision tree-based system for the J-48 algorithm that detects and implements heart disease utilising the two methods of cross-validation and percentage split.
3. The deadliest disease and one of the main causes of death worldwide is heart disease. The medical field is highly dependent on machine learning. In this study, ensemble learning methods are used to improve the accuracy of heart disease prediction. Principal component analysis (PCA) and linear discriminant analysis (LDA) are two characteristics of extraction methods that are used to choose crucial features from the dataset. Selected features are used to compare machine learning algorithms and ensemble learning techniques. Models are evaluated using a variety of techniques, including accuracy, recall, precision, F-measure, and ROC. The outcomes demonstrate that the decision tree-based bagging ensemble learning method produced the best results.