Heart Failure Prediction

Team Members: -

1. Kavana Manvi Krishnamurthy.
2. Sri Varsha Rayabandi
3. Adi Manivarsh

Description: - Cardiovascular diseases (CVDs) are the leading cause of death worldwide, claiming an estimated 17.9 million lives each year, accounting for 31% of global deaths. Heart attacks and strokes are responsible for 80% of CVD deaths, with one-third of these fatalities occurring prematurely in individuals under the age of 70. Heart failure is often a result of CVDs.

Problem Statement: - Early detection and management are crucial for people with cardiovascular conditions or those at high risk due to factors like hypertension, diabetes, hyperlipidemia, or existing disease. Machine learning models can significantly aid in this process.

Dataset Description: - This dataset includes 11 features that can be used to predict the likelihood of heart disease with overall 918 observations(instances).

No of instance (Data quantity): - 918

No of Features: - 11 features (Cardio Vasicular features with Age and sex)

Target: - Heart Disease out put class (1➔heart disease, 0➔ Normal)

Balanced data with 45% of '0' class ad 55% of '1' class.

Pros: -

1. The use-case involves a straightforward task of binary classification, with potential to apply techniques such as Logistic Regression, Navie Bayes and k-NN model.
2. Most of features are generic features to predict heart failure, all the feature makes sense to predict heart failure.
3. No empty/None values.
4. Quantity/Quality.
5. This dataset is mix of numerical and categorical features.
6. Usability score is 10

Cons: -

1. Quantity wise are around 1000 instances, with more instances above 1000 instances can build more accurate models.
2. There can be more other features which are missing and can be more useful to predict heart failure.

Links: -

Kaggle Link: -

[1] Kaggle link: - https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction