**Heart Failure Prediction Dataset**

**SWE 485 AI Project**

**Group 4**

|  |  |
| --- | --- |
| **Name** | **ID** |
| Hams Almalki | 441201320 |
| Alanoud Alomar | 441201113 |
| Leen Alhasan | 439201213 |

**Advisor**

Dr.Nuha Altayash

**Project Motivation:**

we decide to Work on this project because heart failure disease is one of the most common causes of death in the worldwide.

Also, your risk of passing away from a sudden cardiac arrest is increased by the condition.

Five years after being diagnosed with heart failure, 50% of patients will have passed away,

most likely from sudden cardiac arrest.

This makes the illness as deadly as many cancers.

**goal:**

We aim to raise awareness about the risk factors of the heart failure.

People with cardiovascular disease or who are at high cardiovascular risk (due to the presence of one or more risk factors such as hypertension, diabetes, hyperlipidaemia or already established disease) need early detection and management where in a machine learning model can be of great help.

**Source**

This dataset was created by combining different datasets already available independently but not combined before. In this dataset, 5 heart datasets are combined over 11 common features which makes it the largest heart disease dataset available so far for research purposes. The five datasets used for its curation are:

Cleveland: 303 observations

Hungarian: 294 observations

Switzerland: 123 observations

Long Beach VA: 200 observations

Stalog (Heart) Data Set: 270 observations

Total: 1190 observations

Duplicated: 272 observations

Final dataset: 918 observations

The data contains 918 observations with 10 attributes.

**Attribute Information (Column)**

Age: age of the patient [years]

Sex: sex of the patient [M: Male, F: Female]

ChestPainType: chest pain type [TA: Typical Angina, ATA: Atypical Angina, NAP: Non-Anginal Pain, ASY: Asymptomatic]

RestingBP: resting blood pressure [mm Hg]

Cholesterol: serum cholesterol [mm/dl]

FastingBS: fasting blood sugar [1: if FastingBS > 120 mg/dl, 0: otherwise]

RestingECG: resting electrocardiogram results [Normal: Normal, ST: having ST-T wave abnormality (T wave inversions and/or ST elevation or depression of > 0.05 mV), LVH: showing probable or definite left ventricular hypertrophy by Estes' criteria]

MaxHR: maximum heart rate achieved [Numeric value between 60 and 202]

ExerciseAngina: exercise-induced angina [Y: Yes, N: No]

HeartDisease: output class [1: heart disease, 0: Normal]

In these graphs appears the Valid , Mismatch and Missing observation Also, Mean , Standard deviation , Min and Max.



Chart, histogram

Description automatically generated

**Graphical user interface

Description automatically generated with low confidence**

**Graphical user interface, chart, waterfall chart

Description automatically generated with medium confidence**

**Chart, histogram

Description automatically generated**

**Chart

Description automatically generated**

**Link:**

<https://www.kaggle.com/datasets/fedesoriano/heart-failure-prediction?resource=download>