SPRINT-I-PART-1

Date	02 November 2022
Team ID	PNT2022TMID21605
Project Name	Visualizing and Predicting Heart Diseases with an Interactive Dash Board
Maximum Marks	2 Marks

DATASET:

Dataset from Kaggle is collected.

Cardiovascular diseases (CVDs) are the number 1 cause of death globally, taking an estimated 17.9 million lives each year, which accounts for 31% of all deaths worldwide. Four out of 5CVD deaths are due to heart attacks and strokes, and one-third of these deaths occur prematurely in people under 70 years of age. Heart failure is a common event caused by CVDs and this dataset contains 11 features that can be used to predict a possible heart disease.

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 wherein a machine learning model can be of great help.

This data set dates from 1988 and consists of four databases: Cleveland, Hungary, Switzerland, and Long Beach V.

It contains 76 attributes, including the predicted attribute, but all published experiments refer to using a subset of 14 of them. The "target" field refers to the presence of heart disease in the patient. It is integer valued 0 = no disease and 1 = disease.

Attribute Information

1. Age: age of the patient [years]
2. Sex: sex of the patient [M: Male, F: Female]
3. ChestPainType: chest pain type [TA: Typical Angina, ATA: Atypical Angina, NAP: Non-Anginal Pain, ASY: Asymptomatic]
4. RestingBP: resting blood pressure [mm Hg]
5. Cholesterol: serum cholesterol [mm/dl]
6. FastingBS: fasting blood sugar [1: if FastingBS > 120 mg/dl, 0: otherwise]
7. 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]

8. MaxHR: maximum heart rate achieved [Numeric value between 60 and 202]
9. ExerciseAngina: exercise-induced angina [Y: Yes, N: No]
10. Oldpeak: oldpeak = ST [Numeric value measured in depression]
11. ST_Slope: the slope of the peak exercise ST segment [Up: upsloping, Flat: flat, Down: downsloping]
12. HeartDisease: output class [1: heart disease, 0: Normal]

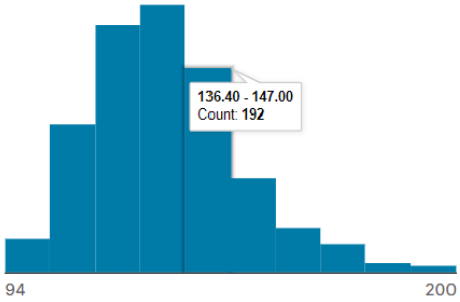
cp

chest pain type

Valid	1025	100%
Mismatched	0	0%
Missing	0	0%
Mean	0.94	
Std. Deviation	1.03	
Quantiles	0	Min
	0	25%
	1	50%
	2	75%
	3	Max

trestbps

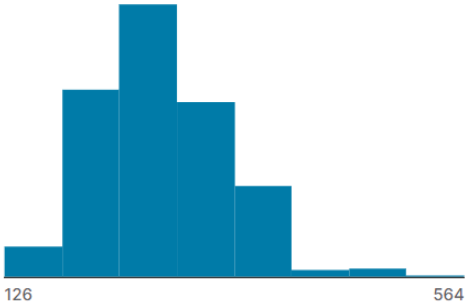
resting blood pressure (in mm Hg on admission to the hospital)



Valid	1025	100%
Mismatched	0	0%
Missing	0	0%
Mean	132	
Std. Deviation	17.5	
Quantiles	94	Min
	120	25%
	130	50%
	140	75%
	200	Max

chol

serum cholestoral in mg/dl



Valid	1025	100%
Mismatched	0	0%
Missing	0	0%
Mean	246	
Std. Deviation	51.6	
Quantiles	126	Min
	211	25%
	240	50%
	275	75%
	564	Max

fbs

(fasting blood sugar > 120 mg/dl) (1 = true; 0 = false)

Valid	1025	100%
Mismatched	0	0%
Missing	0	0%
Mean	0.15	
Std. Deviation	0.36	
Quantiles	0	Min
	0	25%
	0	50%
	0	75%
	1	Max

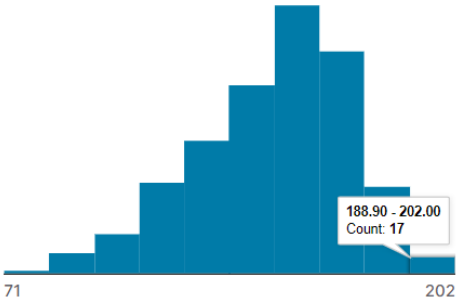
restecg

resting electrocardiographic results

Valid	1025	100%
Mismatched	0	0%
Missing	0	0%
Mean	0.53	
Std. Deviation	0.53	
Quantiles	0	Min
	0	25%
	1	50%
	1	75%
	2	Max

thalach

maximum heart rate achieved



Valid	1025	100%
Mismatched	0	0%
Missing	0	0%
Mean	149	
Std. Deviation	23	
Quantiles	71	Min
	132	25%
	152	50%
	166	75%
	202	Max

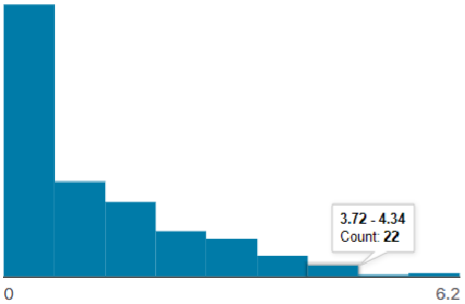
exang

exercise induced angina (1 = yes; 0 = no)

Valid	1025	100%
Mismatched	0	0%
Missing	0	0%
Mean	0.34	
Std. Deviation	0.47	
Quantiles	0	Min
	0	25%
	0	50%
	1	75%
	1	Max

oldpeak

ST depression induced by exercise relative to rest



Valid	1025	100%
Mismatched	0	0%
Missing	0	0%
Mean	1.07	
Std. Deviation	1.17	
Quantiles	0	Min
	0	25%
	0.8	50%
	1.8	75%
	6.2	Max