World Most Deadly Disease

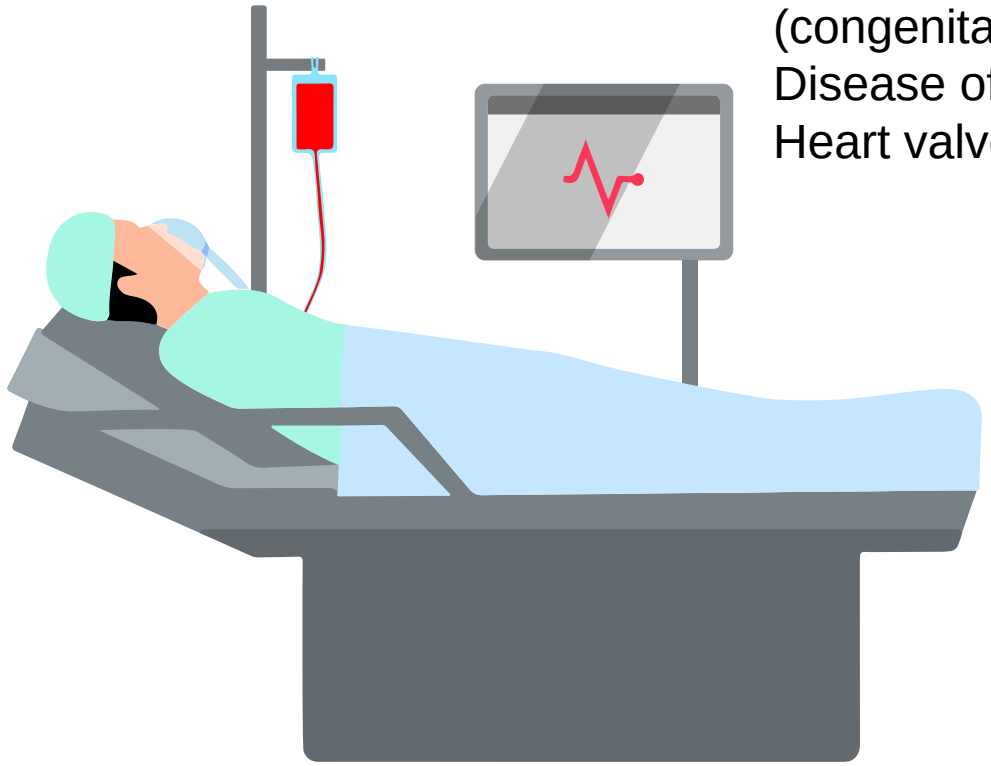
Burhan Rafid Ekatama

Content

Problem Overview
And
Data Understanding

Problem Overview

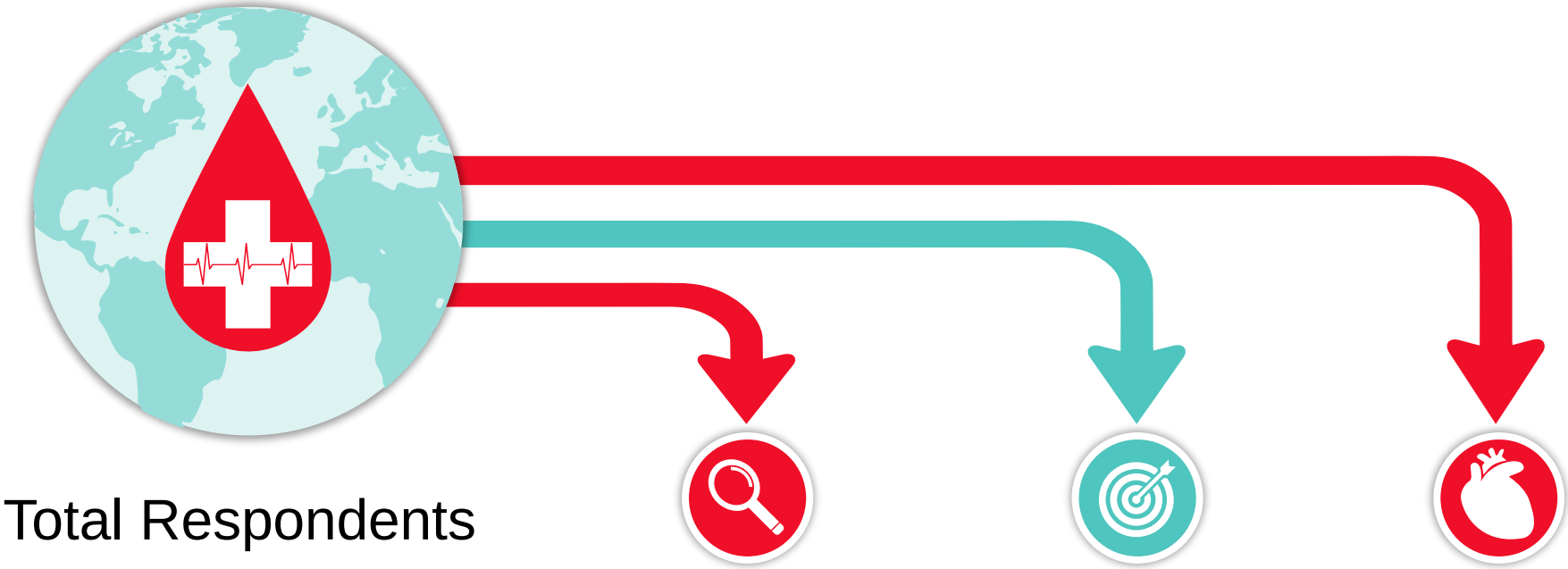
Heart disease describes a range of conditions that affect the heart. Heart diseases include Blood vessel disease, such as coronary artery disease, Irregular heartbeats (arrhythmias), Heart problems born with (congenital heart defects), Disease of the heart muscle, Heart valve disease



Based on cdc.gov, there are number of deaths for leading causes of death

1. **Heart Disease**
2. **Cancer**
3. **COVID - 19**
4. **Accidents**
5. **Stroke**
6. **Chronic Lower Respiratory Disease**

Problem Overview



Total Respondents

70.000 person

Positive and Negative

35021 Person, Negative
34979 Person, Positive

Goals

Predict best model for patients who are suspected of having Heart Disease or not

Hope

Can reduce the risk of death from heart disease and can detect early from the symptoms of the disease

Data Understanding

13 Columns (id, age, gender, height, weight, ap_hi, ap_lo, cholesterol, gluc, smoke, alco, active, cardio)

Missing and duplicated value checking

Columns Relabelling

Delete the row with unreasonable values

Encoding and deleting column “ID”

13 Columns (Age, Height, Weight, Sys_Blood_Press, Dias_Blood_Press, Cholesterol, Glucose, Smoking, Alcoholic, Physical Activity, Cardio Diagnose, Gender_A_Female, Gender_A_Male)

Modelling and Recommendation

Modelling

Modelling Result

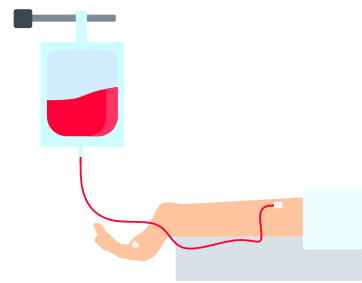
Modelling Performance

	Model	Recall	AUC	F1 Score	Accuracy
0	Logistic_Regression	0.649252	0.707719	0.687864	0.708297
1	Random_Forest	0.694306	0.712908	0.705540	0.713092
2	Decision_Tree	0.633989	0.634045	0.631717	0.634045
3	Extra_Trees	0.692838	0.703541	0.698321	0.703647
4	Gradient_Boosting	0.689169	0.733937	0.719804	0.734380
5	Light_Gradient_Boosting	0.687115	0.733629	0.718980	0.734089
6	Hist_Gradient_Boosting	0.689463	0.734084	0.720000	0.734525

Based on tabel beside, it can be concluded that the best model is **Hist Gradient Boosting**

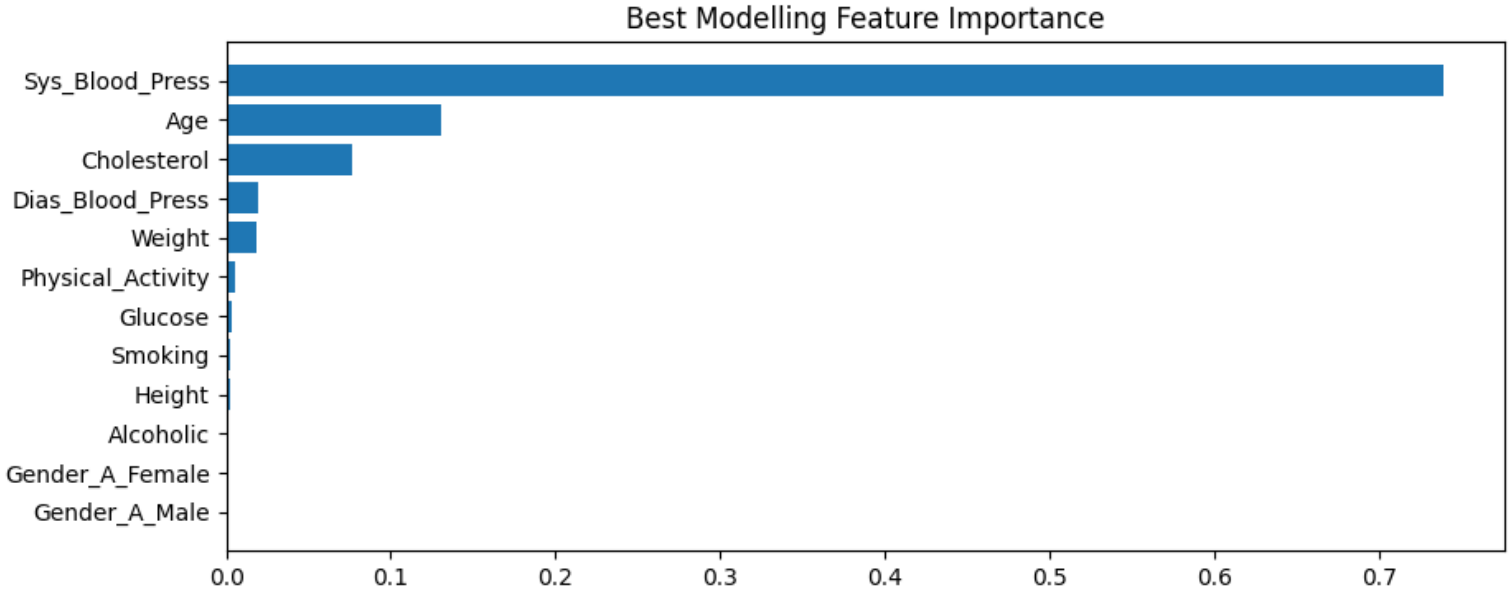
1. Highest **Accuracy** Value (0.7345)
2. Highest **F1 Score** Value (0.72)

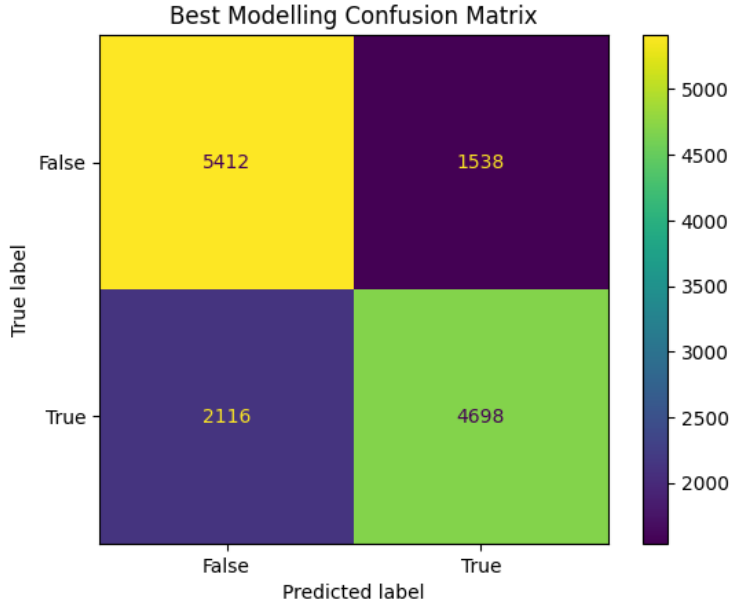
The Recall value on **Hist Gradient Boosting** is **0.689**, meaning that out of 100 person were tested, about **68 – 69 person** get into **Heart Disease**



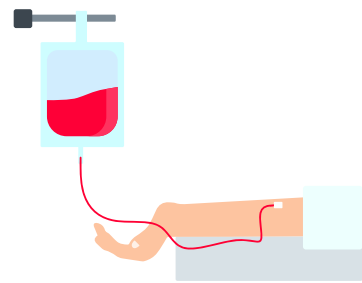
Modelling

Modelling Result





Based on tabel above, it can be concluded that **the most** common cause of **heart disease** is **Systole Blood Pressure**, **Age**, and **Cholesterol rate**



Question
And
Society Insight

Some Questions

Question 1

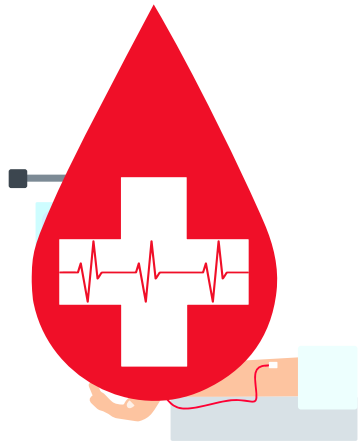
From people with **Heart Disease**, how many on **maximum** and **minimum point** for the **priority**? So we can give more attention into it

Max and Min Value on Age, Weight, Cholesterol, Systole, and Diastole by Positive Respondents

	Category	Max Values	Min Values
0	Age (year)	64.970000	39.110000
1	Weight	200.000000	21.000000
2	Cholesterol	3.000000	1.000000
3	Systole	240.000000	70.000000
4	Diastole	190.000000	8.000000

From table beside, we know that all of positive respondent were in middle-aged, and a wide range of weight, Systole rate, and Diastole rate.

Everyone has the possibility of this disease



Some Questions

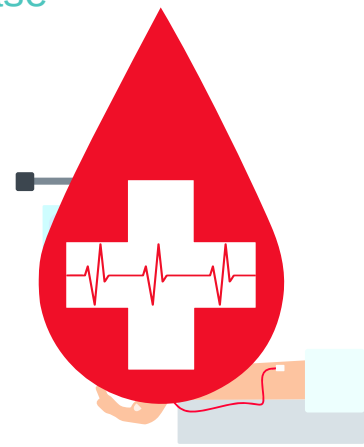
Question 2

Separated by **cholesterol** (most common cause of narrowing of the arteries), how **mean** of the **Age** and **Systole**?

Cholesterol	Cholesterol_M	Age	Sys_Blood_Press
1	Normal	19954.34	132.47
2	Below_Normal	19878.73	137.56
3	High	20628.66	135.36

From table beside, from all of positive Heart Disease respondents we know that on **normal cholesterol rate**, has the mean age is 19954,34 days (**54+ years old**) and mean Systole is **132.47**. On **high cholesterol rate**, has the mean age is 20628.66 days (**56+ years old**) and mean Systole is **135.36**

So we can conclude that person is 54 years old and above, with high Systole (standard value is 120) has a **high risk** of heart disease



Society Insight

Heart disease describes a range of **conditions that affect the heart**. Heart diseases include:

- Blood vessel disease, such as coronary artery disease
- Irregular heartbeats (arrhythmias)
- Heart problems you're born with (congenital heart defects)
- Disease of the heart muscle
- Heart valve disease

Many forms of heart disease can be prevented or treated with **healthy lifestyle choices**.

mayoclinic.org

Society Insight

Coronary artery disease is a common heart condition that affects the major blood vessels that supply the heart muscle. **Cholesterol deposits (plaques)** in the heart arteries are usually the cause of **coronary artery disease**.

Symptoms of coronary artery disease can include chest pain, chest tightness, chest pressure and chest discomfort (angina), shortness of breath, pain in the throat, upper belly area or back, weakness or coldness in the legs or arms if the blood vessels in those body areas are narrowed

Risk factors for heart disease include:

Age, sex, family history, smoking, unhealthy diet, high blood pressure, high cholesterol, diabetes, obesity, lack of exercise, stress

mayoclinic.org

Society Insight

There is some **tips** which you can use to **prevention** the heart disease

- Don't smoke.
- Eat a diet that's low in salt and saturated fat.
- Exercise at least 30 minutes a day on most days of the week.
- Maintain a healthy weight.
- Reduce and manage stress.
- Control high blood pressure, high cholesterol and diabetes.
- Get good sleep. Adults should aim for 7 to 9 hours daily.

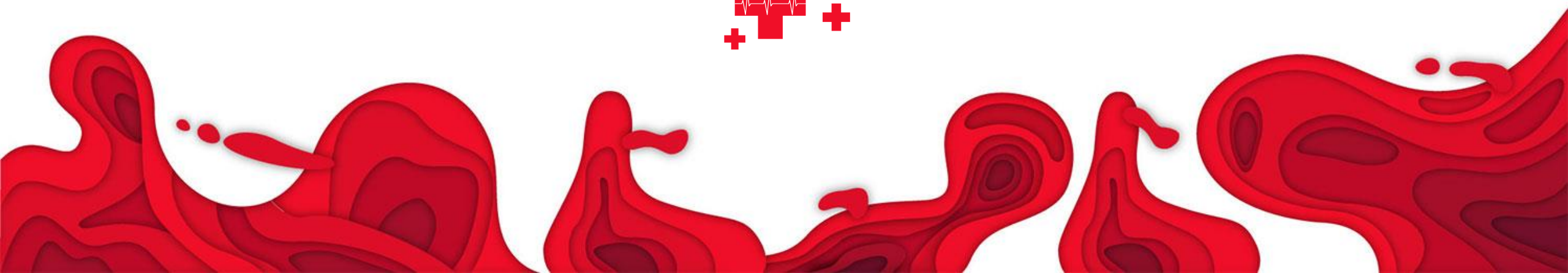
Based on the best model chosen, it is hoped that it can more **quickly detect** a person or patient who is suspected of having a Heart Disease. Because the **faster** the first treatment of a patient can **increase the percentage of recovery** from that patient.

mayoclinic.org

References

https://www.cdc.gov/nchs/fastats/leading-causes-of-death.html

https://www.mayoclinic.org/diseases-conditions/heart-disease/symptoms-causes/syc-20353118



THANK YOU

Reach me anytime:

Project link: