

Heart Disease Detector

Ayush Jadhav – June 2025

Problem Statement

- Heart disease is a leading cause of death globally.
- 18 million people worldwide die every year due undiagnosed heart diseases
- Early detection is key to timely treatment and prevention.
- Patients have to wait for weeks to get an appointment.
- Manual diagnosis can be slow and inconsistent.

Goal :

Build a machine learning model that predicts whether a patient has heart disease based on clinical data.

Dataset & Features

- **Total Records:** 1190 patients
- **Total Features:** 11
- **Target Variable:** 1 = Disease, 0 = No Disease

#	A	B	C	D	E	F	G	H	I	J	K	L
1	age	sex	chest pain type	resting bp s	cholesterol	fasting blood sugar	resting ecg	max heart rate	exercise angina	oldpeak	ST slope	target
2	40	1	2	140	289	0	0	172	0	0	1	0
3	49	0	3	160	180	0	0	156	0	1	2	1
4	37	1	2	130	283	0	1	98	0	0	1	0
5	48	0	4	138	214	0	0	108	1	1.5	2	1
6	54	1	3	150	195	0	0	122	0	0	1	0
7	39	1	3	120	339	0	0	170	0	0	1	0
8	45	0	2	130	237	0	0	170	0	0	1	0
9	54	1	2	110	208	0	0	142	0	0	1	0
10	37	1	4	140	207	0	0	130	1	1.5	2	1
11	48	0	2	120	284	0	0	120	0	0	1	0
12	37	0	3	130	211	0	0	142	0	0	1	0
13	58	1	2	136	164	0	1	99	1	2	2	1
14	39	1	2	120	204	0	0	145	0	0	1	0
15	49	1	4	140	234	0	0	140	1	1	2	1

Attribute	Code given	Unit	Data type
age	Age	in years	Numeric
sex	Sex	0 = female, 1 = male	Binary
chest pain type	chest pain type	1 = typical angina, 2 = atypical angina, 3 = non-anginal pain, 4 = asymptomatic	Nominal
resting blood pressure	resting bp s	in mm Hg	Numeric
serum cholesterol	cholesterol	in mg/dl	Numeric
fasting blood sugar	fasting blood sugar	1 = sugar > 120mg/dL 0 = sugar < 120mg/dL	Binary
resting electrocardiogram results	resting ecg	0 = normal, 1 = ST-T wave abnormality (T wave inversions and/or ST elevation/depression of > 0.05 mV), 2 = Probable or Definite Left Ventricular hypertrophy by Estes' criteria	Nominal
maximum heart rate achieved	max heart rate	71–202	Numeric
exercise induced angina	exercise angina	0 = no, 1 = yes	Binary
oldpeak =ST	oldpeak	depression	Numeric
the slope of the peak exercise ST segment	ST slope	1 = upward 2 = flat, 3 = downward	Nominal
class	target	0 = Normal, 1 = Heart Disease	Binary

Model & Approach

➤ Models Tested

- GaussianNB
- Random Forest Classifier
- Gradient Boosting Classifier
- Logistic Regression
- KNeighbors Classifier

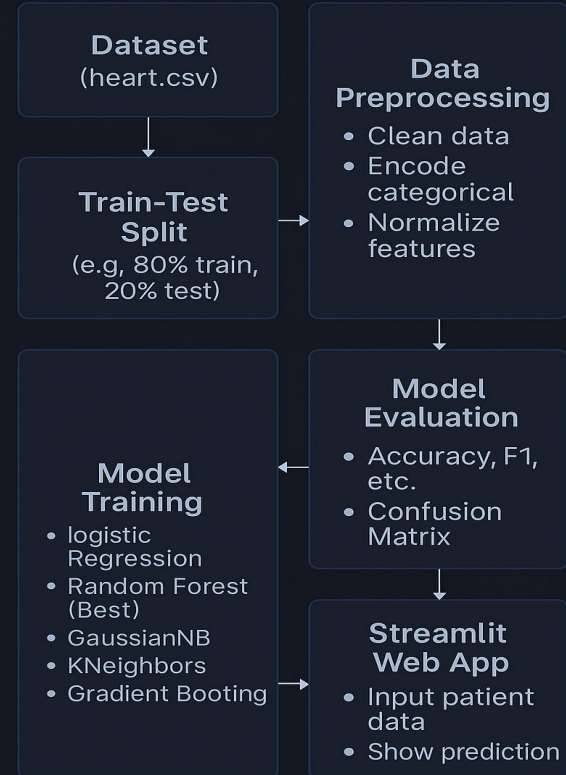
➤ Data Preprocessing

- Clean Dataset (no missing value)
- Scaled numeric features (standardization)
- Split into train/test (80/20)

➤ Tools used

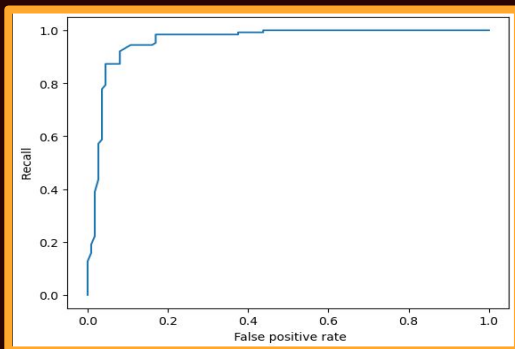
- Python 3.8
- Jupyter Notebook
- Scikit-learn
- Pandas
- Pickle
- Streamlit

Heart Disease Detection Pipeline

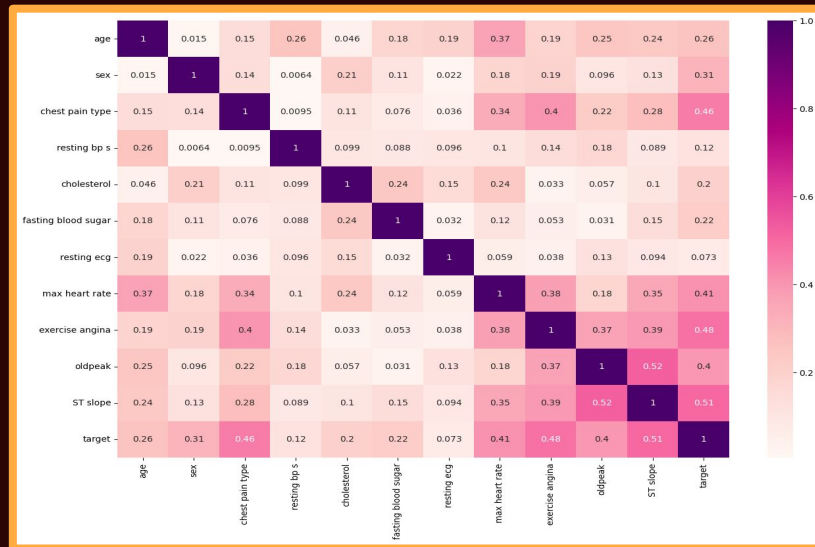


Results & Evaluation

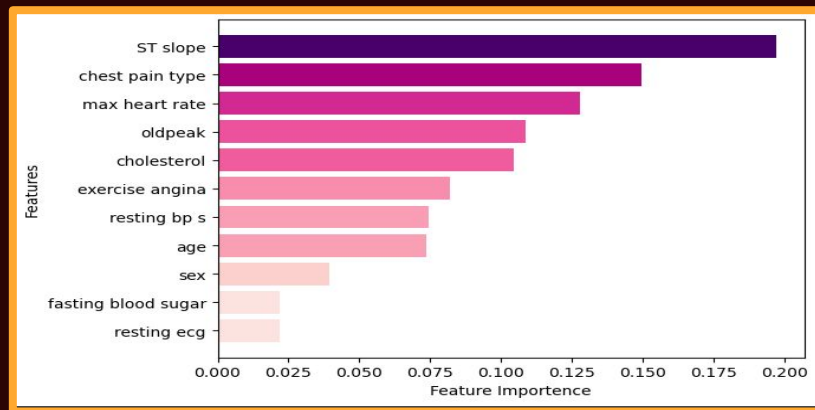
- **Best Accuracy: ~91% (Random Forest)**
- **Recall Score: ~87%**
- **Metrics used:**
 - Accuracy, Precision, Recall, F1-Score
 - Confusion Matrix
- **Most Important features:**
 - ST slope
 - Chest pain type
 - Max heart rate
- **Well-balanced performance across classes**



ROC CURVE



Confusion Matrix



Feature Analysis

Streamlit App & Conclusion

- Built a real-time web app using **Streamlit**
 - User can input values and get a prediction instantly
 - Simple interface designed for usability
-
- **Conclusion:**
 - ML can support doctors in faster diagnosis
 - Future work: Improve model, use larger datasets, add explainability

Get in touch:

contact.ayushjadhav@gmail.com

[Linkdin](#)

Want to try out the app?

[Click here for a live demo](#)

Source code:

[Github Repository](#)

System to predict Heart Disease

Heart Disease Prediction App

Enter patient data to predict the risk of heart disease

Patient Name

John Doe

Age

45

Sex

Male

Chest Pain Type

0

Resting Blood Pressure

120

Cholesterol

200

Fasting Blood Sugar > 120 mg/dl

0

Resting ECG

0

Max Heart Rate Achieved

150

Exercise Induced Angina

0

Oldpeak (ST depression)

1.00

ST Slope

0

Predict

John Doe is Unlikely to have heart disease