

# HEART FAILURE PREDICTION ANALYTICS

**Advances in Data Science / Architecture[INFO 7370]**  
**Professor Ram Kumar Harihar**

Team **Data Driven:**

Akshita Barot (002704943)

Gauri Jaydeep Patole (002700996)

Hrishikesh Warriar (002768017)

Parvathy Pillai (002775764)

### **Global Impact:**

12M annual Heart Disease deaths (WHO)

Rising cardiovascular cases worldwide

### **Research and Danger:**

Research identifies risks and predicts outcomes

Heart Disease often silent but deadly

### **Early Detection Matters:**

Early diagnosis aids lifestyle adjustments

Reduces heart disease complications

### **Project Focus:**

Predicting Heart Disease via ML

ML aids timely detection and management

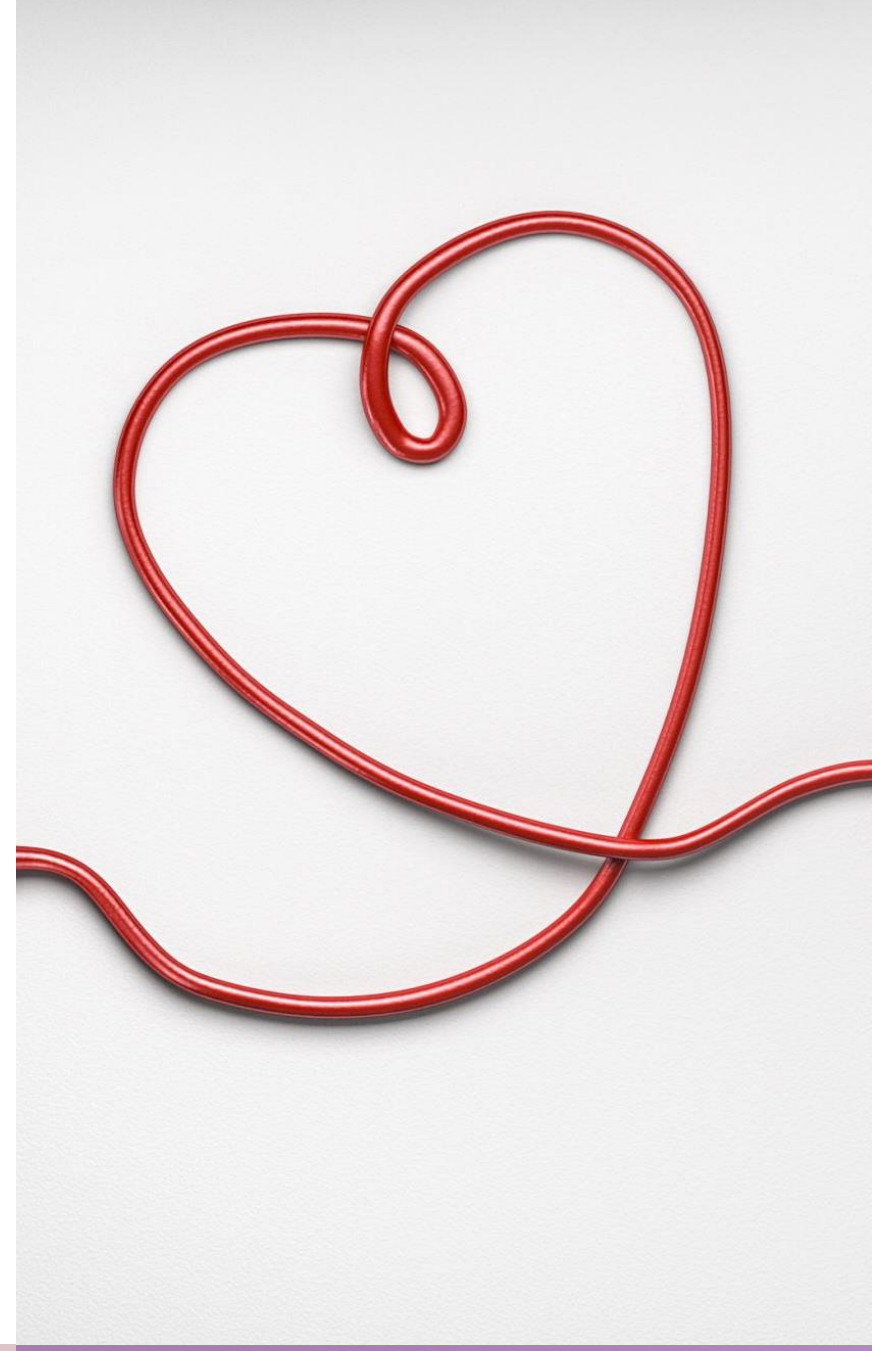
### **Data-Driven Insights:**

Analyzing patient data for risk trends

Various data techniques used

## → INTRODUCTION

- - Heart diseases are the leading global cause of death, accounting for over a million deaths annually.
- - Approximately one third of these deaths occur before the age of 70, highlighting the significance of premature mortality.
- - Researchers worldwide are dedicated to preventing, managing, and ultimately finding a cure for heart diseases.
- - High cardiovascular risk is associated with factors like hypertension, diabetes, hyperlipidemia, and established diseases.
- - Early detection and management are crucial for individuals with cardiovascular disease or high risk.
- - Supervised Machine learning models offer potential assistance in predicting heart failure risk.



## → DATASET FEATURES

- **Age:** age of the patient [years]
- **Sex:** sex of patient [M:Male, F:Female]
- **ChestPainType:** chest pain type[TA: typical Angina, ATA: Atypical Angina, NAP: No 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), LV 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]
- **Oldpeak:** oldpeak = ST [Numeric value measured in depression];
- **ST\_Slope:** the slope of the peak exercise ST segment [Up: upsloping, Flat : flat, Down: downsloping]
- **HeartDisease :** output class [1: heart disease, 0: Normal]



DATASET

## → ALGORITHMS USED

Decision Tree

Random Forest

MLP Classifier

SVM

Extra Tree

Logistic Regression

KNN

Gradient Boosting Classifier

Stacking Classifier

→ Best Performing Model: Random Forest

Model	Train_accuracy	Test_accuracy	Precision_score	Recall_score	F1_score
Decision Tree	0.897196	0.811594	0.838926	0.816993	0.827815
Random Forest	0.925234	0.851449	0.863636	0.869281	0.866450
MLP	0.894081	0.858696	0.870130	0.875817	0.872964
SVC	0.898754	0.855072	0.864516	0.875817	0.8070130
Extra Tree	0.900312	0.847826	0.858065	0.869281	0.863636
Gradient Boosting	0.914330	0.851449	0.873333	0.856209	0.864686
Logistic Regression	0.853583	0.815217	0.844595	0.816993	0.830565
KNN	0.878505	0.829710	0.848684	0.843137	0.845902
Stacking	0.934579	0.844203	0.871622	0.843137	0.857143



## → CONCLUSION

- From the employed list of algorithms, Random Forest and additionally Stacking Classifier showed the top performance with the following performance metrics, Random Forest -
- Train Accuracy: 92.52% | Test Accuracy: 85.814% | f1 Score: 86.64%
- Stacking Classifier -
- Train Accuracy: 93.45% | Test Accuracy: 84.42% | f1 Score: 85.71%

