HEART DISEASE ANALYSIS

Data Analytics & Tableau Project Report

Project Type: Data Analytics with Machine Learning & Tableau Dashboard

Dataset Size: 1000 Patient Records

Tools Used: Python, Pandas, Scikit-learn, Flask, Tableau, Bootstrap

1. Introduction

This project focuses on analyzing heart disease patient data using Data Analytics and Tableau. The system evaluates 1000 patient records to identify patterns, risk factors, and predictive indicators of heart disease.

2. Objectives

• Analyze patient health data.
• Identify key risk factors.
• Build machine learning models for prediction.
• Create interactive Tableau dashboards.
• Deploy prediction system using Flask.

3. Existing System

Traditional health analysis systems rely on manual diagnosis and static reports without real-time analytics or visualization.

4. Disadvantages of Existing System

• Manual interpretation required.
• No predictive analysis.
• Limited visualization.
• No real-time dashboard.

5. Proposed System

The proposed system uses Data Analytics and Machine Learning to analyze heart disease data, visualize patterns through Tableau dashboards, and provide predictive insights via a Flask web application.

6. System Architecture

Data Collection → Data Preprocessing → Exploratory Data Analysis → Model Training → Model Evaluation → Tableau Dashboard → Flask Web Deployment.

7. Dataset Description

The dataset contains 1000 patient records with 13 medical features including age, sex, cholesterol, blood pressure, max heart rate, chest pain type, fasting blood sugar, and others.

8. Data Preprocessing

• Handling missing values.
• Feature scaling.
• Encoding categorical variables.
• Splitting into training and testing datasets.

9. Machine Learning Models Used

• Logistic Regression
• Random Forest
• Gradient Boosting
• Support Vector Machine (SVM)

10. Model Performance

Best Accuracy: 72%
Best AUC: 64%
Cross Validation Accuracy: ~71%

11. Tableau Dashboard

Interactive dashboards include:
• Disease distribution
• Age group analysis
• Correlation heatmap
• Chest pain risk analysis
• Feature importance visualization

12. Advantages

• Interactive dashboards
• Data-driven decision making
• Predictive analysis
• Web-based deployment

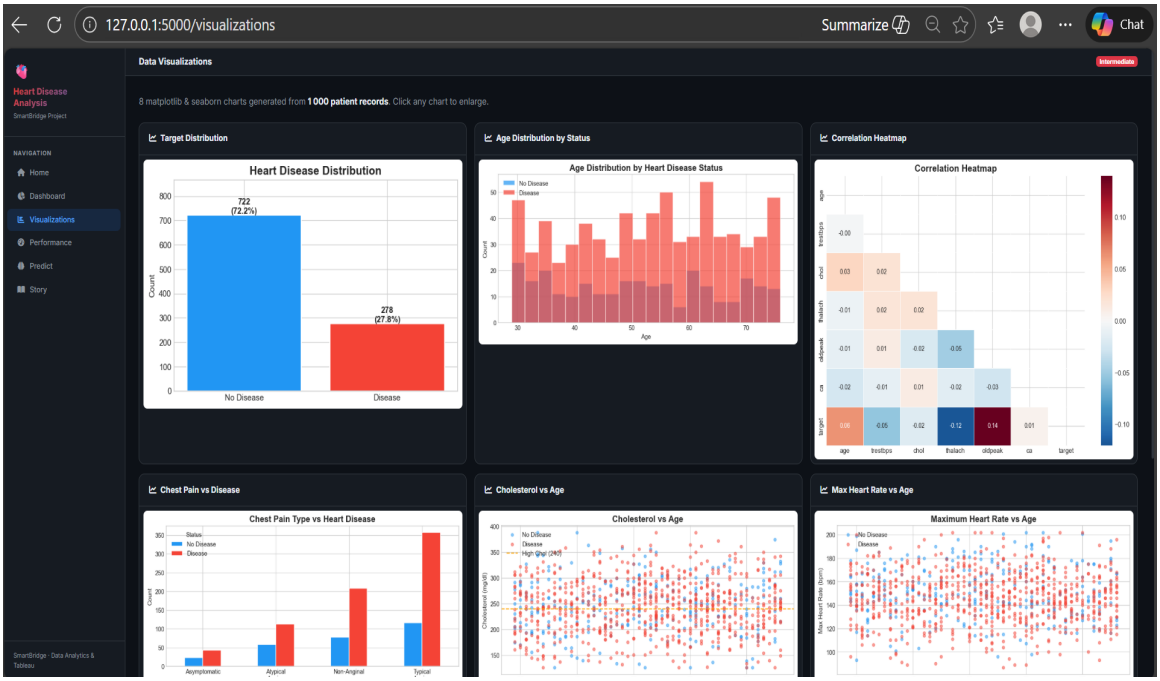
13. Conclusion

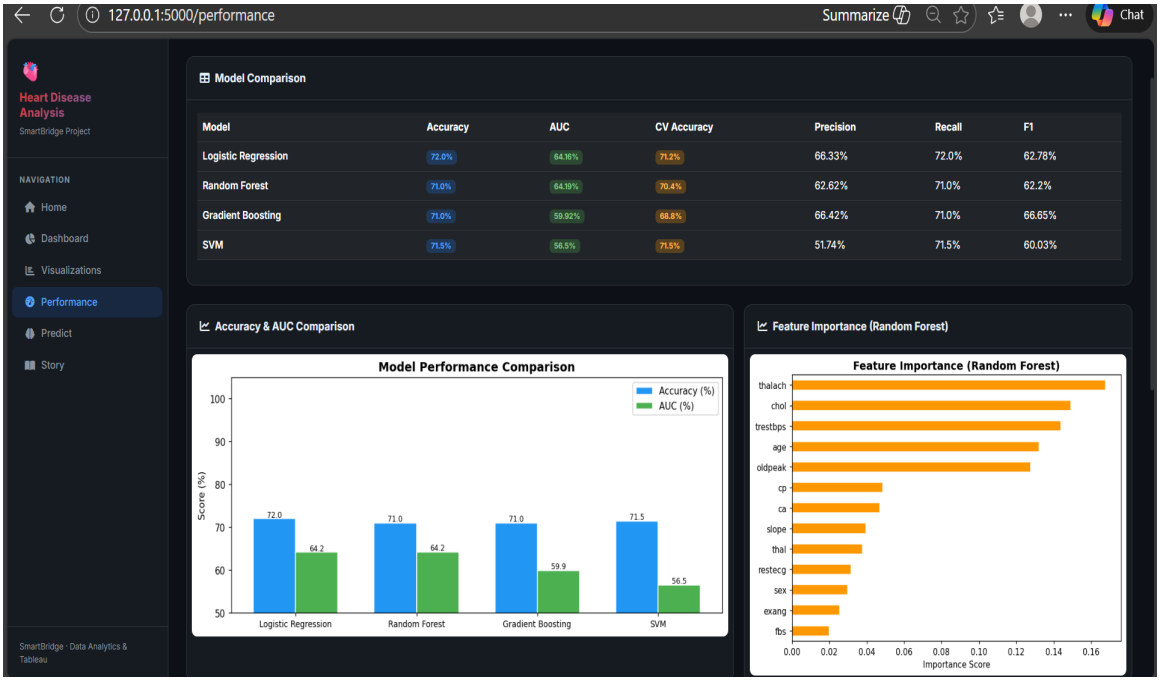
The project successfully demonstrates how Data Analytics and Tableau can be integrated with Machine Learning to predict heart disease and visualize patient health trends.

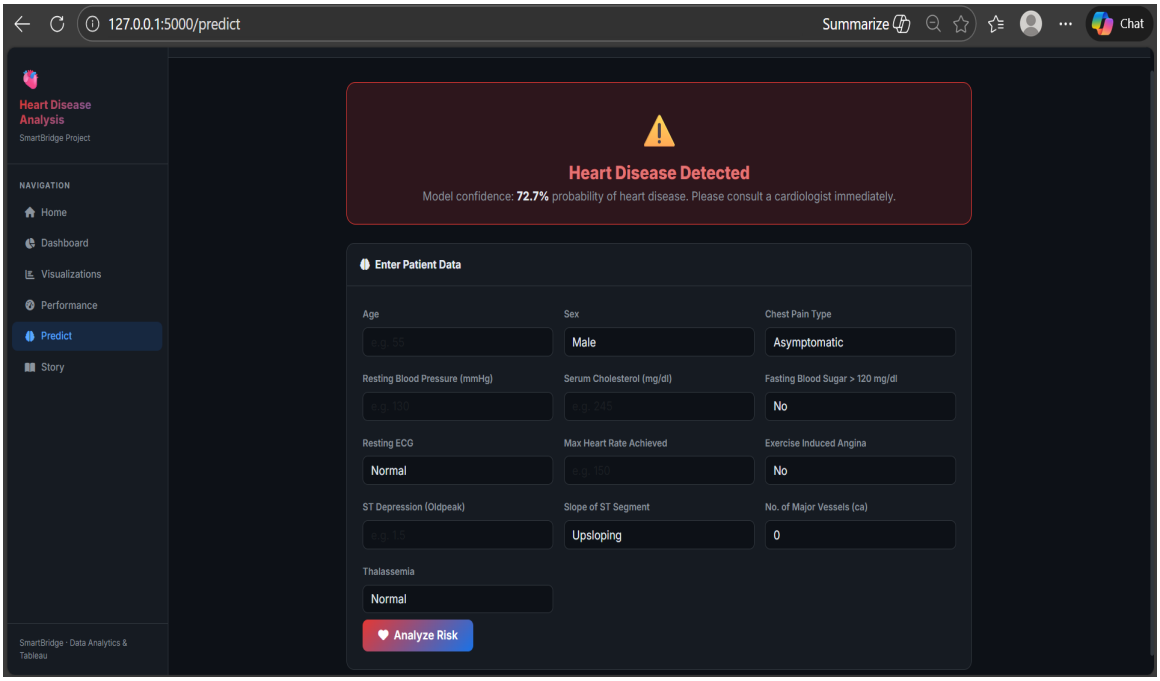
14. Future Scope

• Real-time hospital integration
• Cloud deployment
• Deep learning models
• Integration with IoT health devices

Project Screenshots


127.0.0.1:5000/visualizations
Summarize
Chat
Heart Disease Analysis
SmartBridge Project
NAVIGATION
Home
Dashboard
Visualizations
Performance
Predict
Story
Data Visualizations
Intermediate
8 matplotlib & seaborn charts generated from 1 000 patient records. Click any chart to enlarge.
Target Distribution
Heart Disease Distribution
722 (72.2%)
278 (27.8%)
No Disease
Disease
Count
Age Distribution by Status
Age Distribution by Heart Disease Status
No Disease
Disease
Age
Correlation Heatmap
Chest Pain vs Disease
Chest Pain Type vs Heart Disease
Asymptomatic
Atypical
Non-Anginal
Typical
Cholesterol vs Age
Max Heart Rate vs Age
Maximum Heart Rate vs Age
SmartBridge · Data Analytics & Tableau


127.0.0.1:5000/performance
Summarize
Chat
Heart Disease Analysis
SmartBridge Project
NAVIGATION
Home
Dashboard
Visualizations
Performance
Predict
Story
Model Comparison
Model | Accuracy | AUC | CV Accuracy | Precision | Recall | F1
Logistic Regression | 72.0% | 64.16% | 71.2% | 66.33% | 72.0% | 62.78%
Random Forest | 71.0% | 64.19% | 70.4% | 62.62% | 71.0% | 62.2%
Gradient Boosting | 71.0% | 59.92% | 68.8% | 66.42% | 71.0% | 66.65%
SVM | 71.5% | 56.5% | 71.5% | 51.74% | 71.5% | 60.03%
Accuracy & AUC Comparison
Model Performance Comparison
Accuracy (%)
AUC (%)
Score (%)
72.0
64.2
71.0
64.2
71.0
59.9
71.5
56.5
Logistic Regression
Random Forest
Gradient Boosting
SVM
Feature Importance (Random Forest)
Feature Importance (Random Forest)
thalach
chol
trestbps
age
oldpeak
cp
ca
slope
thal
restecg
sex
exang
fbs
Importance Score
SmartBridge · Data Analytics & Tableau


127.0.0.1:5000/predict
Summarize
Chat
Heart Disease Analysis
SmartBridge Project
NAVIGATION
Home
Dashboard
Visualizations
Performance
Predict
Story
SmartBridge · Data Analytics & Tableau
Heart Disease Detected
Model confidence: 72.7% probability of heart disease. Please consult a cardiologist immediately.
Enter Patient Data
Age
e.g. 55
Sex
Male
Chest Pain Type
Asymptomatic
Resting Blood Pressure (mmHg)
e.g. 130
Serum Cholesterol (mg/dl)
e.g. 245
Fasting Blood Sugar > 120 mg/dl
No
Resting ECG
Normal
Max Heart Rate Achieved
e.g. 150
Exercise Induced Angina
No
ST Depression (Oldpeak)
e.g. 1.0
Slope of ST Segment
Upsloping
No. of Major Vessels (ca)
0
Thalassemia
Normal
Analyze Risk