Data Analytics & Tableau Project Documentation

Project Title: Heart Disease Analysis

Team ID: LTVIP2026TMIDS64979

Team Size: 5

Team Members:

1. Dhannodi Hemanth Kumar (Team Leader)
2. Dileep Tunga
3. Karnati Nagendra Prasad
4. Chinthakayala Balaji
5. Dharmavarapu Pavan Kumar

1. Introduction

Heart Disease Analysis is a Data Analytics and Machine Learning project that analyzes 1000 patient records to identify patterns and predict the probability of heart disease. The project integrates Python, Scikit-learn, Tableau dashboards, and Flask web deployment.

2. Project Overview

Goal: To develop an intelligent heart disease prediction and visualization system.

1. Interactive Dashboard with disease statistics
2. Exploratory Data Analysis using charts & heatmaps
3. Machine Learning Models (Logistic Regression, Random Forest, SVM, Gradient Boosting)
4. Model Performance Comparison
5. Flask-based prediction web interface
6. Feature Importance Analysis

3. Architecture

Frontend: HTML, Bootstrap, Flask Templates Backend: Python (Flask) Machine Learning: Scikit-learn Visualization: Matplotlib, Seaborn, Tableau Dataset: 1000 patient medical records

4. Setup Instructions

1. Install Python 3.8+
2. Install required libraries using pip install -r requirements.txt
3. Run Flask application using python app.py
4. Access project at http://127.0.0.1:5000

5. Folder Structure

1. app.py (Main Flask Application)
2. templates/ (HTML Pages)
3. static/ (CSS, JS, Images)
4. models/ (Saved ML Models)
5. dataset/ (Heart Disease CSV Data)

6. API Endpoints

1. GET / → Home Page
2. GET /dashboard → Dashboard Overview
3. GET /visualizations → Data Charts
4. GET /performance → Model Comparison
5. GET /predict → Prediction Form
6. POST /predict → Heart Disease Prediction Result

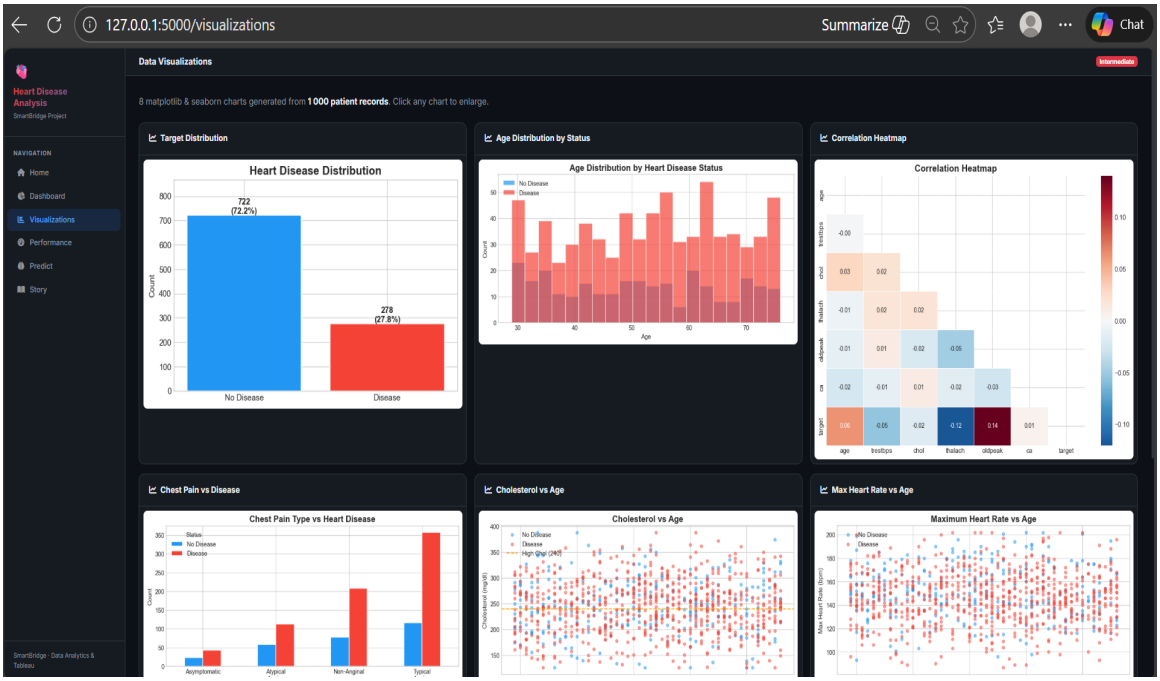
7. User Interface Screens

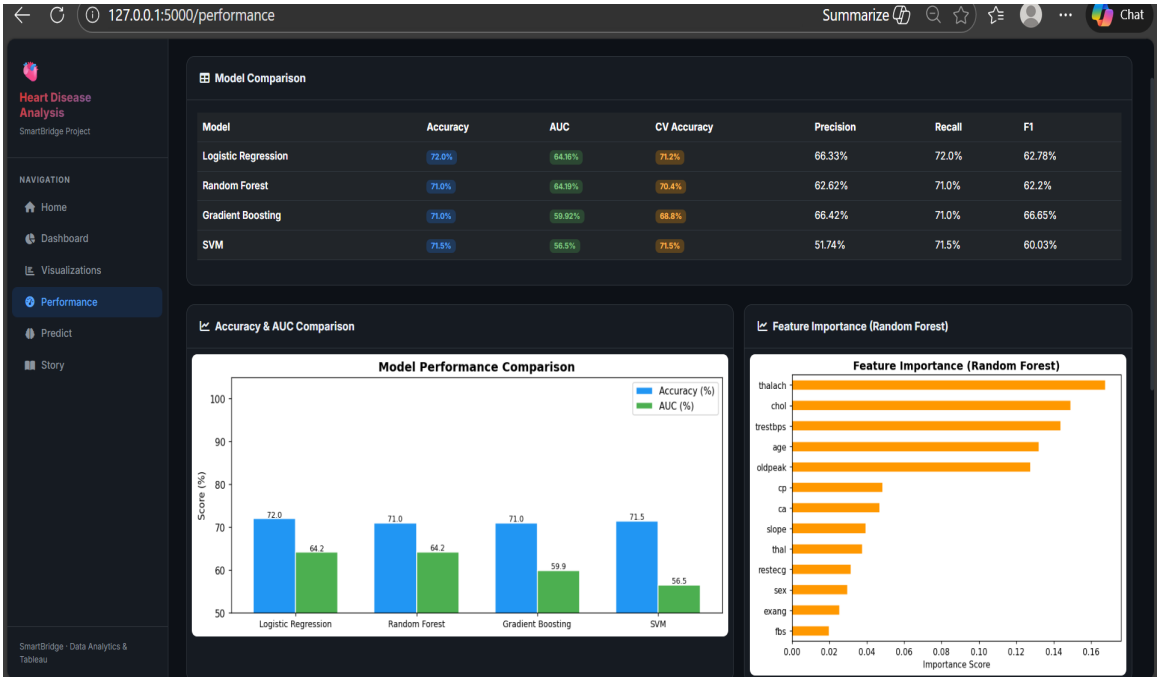
1. Home Page (Project Overview)
2. Dashboard (Disease Statistics & Charts)
3. Visualizations Page (EDA Graphs)
4. Performance Page (Model Comparison & Feature Importance)
5. Prediction Page (Patient Data Input Form)

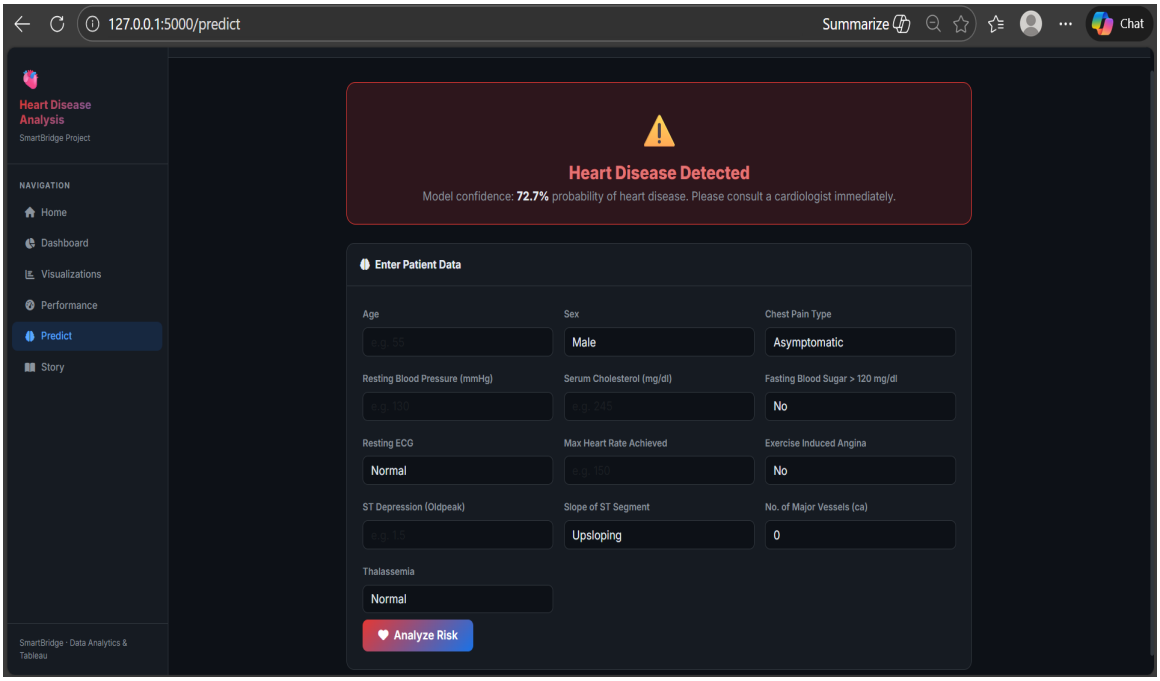
8. Testing

Frontend Testing: UI responsiveness and navigation tested in Chrome. Backend Testing: Model accuracy, API response validation. Model Validation: Cross-validation and confusion matrix evaluation.

9. 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