HEART DISEASE PREDICTION SYSTEM USING MACHINE LEARNING

Name: SURAJIT GIRI
Group Members: SURAJIT GIRI & SURAJIT DAS
College: Haldia Institute of Technology
Internship Program: Advanced Python Internship
Organization: Euphoria GenX
Submission Date: 20/02/2026

ABSTRACT

This project presents a Machine Learning-based Heart Disease Prediction System designed to predict the likelihood of heart disease using clinical parameters. Logistic Regression was used for classification. The model achieved 87% accuracy and was deployed using Streamlit for user interaction.

INTRODUCTION

Heart disease is one of the leading causes of mortality worldwide. Early detection can significantly reduce risks. This project builds a classification model to predict heart disease based on medical input features.

OBJECTIVES

• Analyze healthcare data
• Build and train ML model
• Evaluate performance
• Deploy web application

TECHNOLOGIES USED

Python, Pandas, NumPy, Scikit-learn, Matplotlib, Streamlit

DATASET DESCRIPTION

Dataset contains 303 patient records with 13 input features including age, sex, chest pain type, blood pressure, cholesterol, heart rate etc. Target variable indicates presence (1) or absence (0) of heart disease. Data split: 80% training, 20% testing.

METHODOLOGY

Data collection → Preprocessing → Train-test split → Logistic Regression model training → Evaluation → Deployment using Streamlit.

PERFORMANCE METRICS

Accuracy: 87%
Precision: 85%
Recall: 86%
F1 Score: 85%

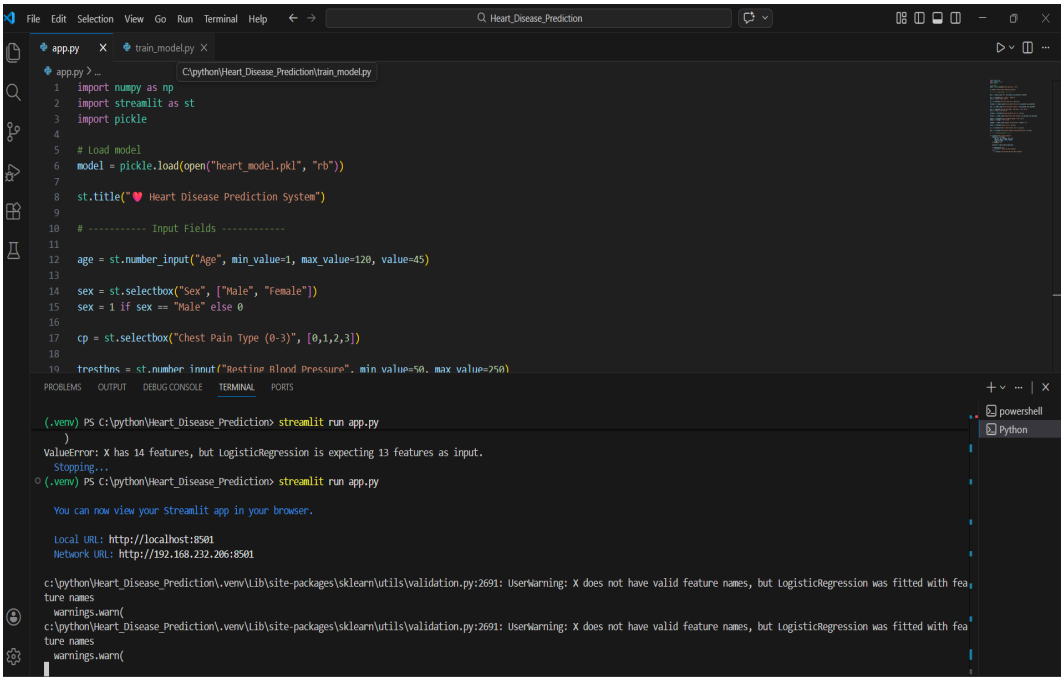
CONCLUSION

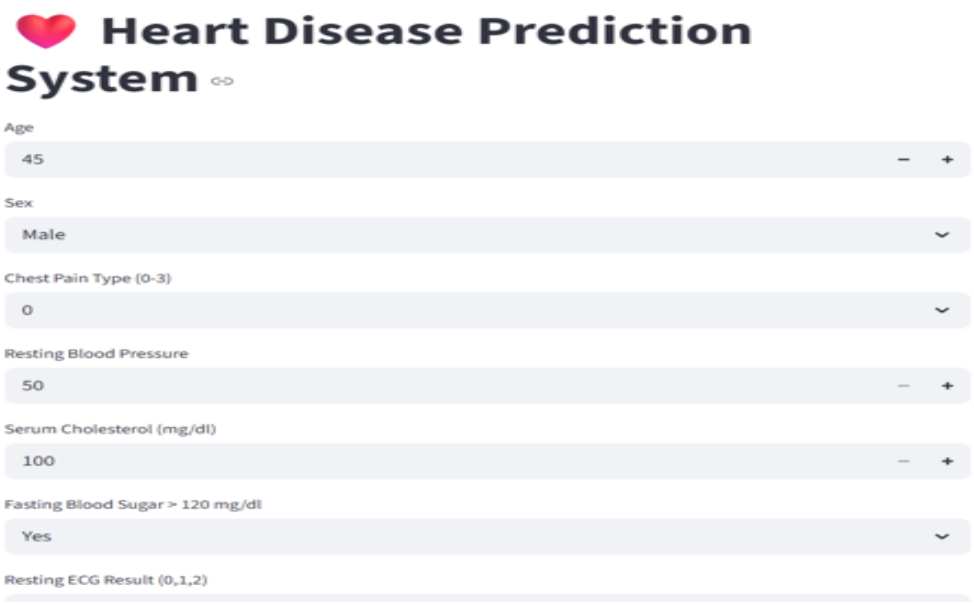
The system successfully demonstrates the use of machine learning in healthcare prediction. It should be used as a supportive tool, not as a replacement for medical professionals.

FUTURE SCOPE

Implement advanced algorithms, improve dataset size, deploy on cloud platforms, develop mobile application.

PROJECT SCREENSHOTS


File Edit Selection View Go Run Terminal Help
Heart_Disease_Prediction
app.py
train_model.py
C:\python\Heart_Disease_Prediction\train_model.py
import numpy as np
import streamlit as st
import pickle
# Load model
model = pickle.load(open("heart_model.pkl", "rb"))
st.title("❤️ Heart Disease Prediction System")
# ---------- Input Fields ----------
age = st.number_input("Age", min_value=1, max_value=120, value=45)
sex = st.selectbox("Sex", ["Male", "Female"])
sex = 1 if sex == "Male" else 0
cp = st.selectbox("Chest Pain Type (0-3)", [0,1,2,3])
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
(.venv) PS C:\python\Heart_Disease_Prediction> streamlit run app.py
)
ValueError: X has 14 features, but LogisticRegression is expecting 13 features as input.
Stopping...
(.venv) PS C:\python\Heart_Disease_Prediction> streamlit run app.py
You can now view your Streamlit app in your browser.
Local URL: http://localhost:8501
Network URL: http://192.168.232.206:8501
c:\python\Heart_Disease_Prediction\.venv\Lib\site-packages\sklearn\utils\validation.py:2691: UserWarning: X does not have valid feature names, but LogisticRegression was fitted with fea
ture names
warnings.warn(
c:\python\Heart_Disease_Prediction\.venv\Lib\site-packages\sklearn\utils\validation.py:2691: UserWarning: X does not have valid feature names, but LogisticRegression was fitted with fea
ture names
warnings.warn(
powershell
Python


❤️ Heart Disease Prediction System
Age
45
Sex
Male
Chest Pain Type (0-3)
0
Resting Blood Pressure
50
Serum Cholesterol (mg/dl)
100
Fasting Blood Sugar > 120 mg/dl
Yes
Resting ECG Result (0,1,2)

0

Maximum Heart Rate Achieved

50

Exercise Induced Angina

Yes

Oldpeak (ST depression)

0.0

Slope (0,1,2)

0

Number of Major Vessels (0-3)

0

Thalassemia (0=Normal,1=Fixed,2=Reversible)

2

Predict Heart Disease

Person does NOT have Heart Disease