

What is Your Heart Rate Telling You?

Heart Disease Prediction ML Model

Model Summary

Algorithm Used: Logistic Regression

Library Used:

- `scikit-learn (sklearn)`
- `pandas` # Data handling
- `numpy` # Numerical operations
- `from sklearn.model_selection import train_test_split`
- `from sklearn.linear_model import LogisticRegression`
- `from sklearn.preprocessing import StandardScaler, LabelEncoder`
- `from sklearn.metrics import accuracy_score, confusion_matrix, roc_curve, roc_auc_score, ConfusionMatrixDisplay`

Model File: model.pkl

Scaler File: scaler.pkl

Target Variable: HeartDisease (1 = Disease, 0 = No Disease)

Features Used for Prediction:

- Age
- Sex (M/F encoded as 1/0)
- ChestPainType (encoded 0–3)
- Cholesterol
- MaxHR (Maximum Heart Rate)
- ExerciseAngina (Yes/No encoded as 1/0)

What is Logistic Regression?

Logistic Regression is a supervised learning classification algorithm.

Even though it has “regression” in its name, it’s used for binary classification tasks — like predicting if a patient has heart disease (1) or not (0).

It works by calculating a weighted sum of the input features, and applying a sigmoid function to output a value between 0 and 1 — interpreted as the probability of the positive class.

Model Training Flow:

Data Preprocessing

Categorical variables like **Sex, ChestPainType, and ExerciseAngina** were encoded.

Feature scaling was applied using StandardScaler.

Model Training :

The data was split into 80% train and 20% test.

A LogisticRegression model was trained using the scaled data.

Model Evaluation :

Accuracy was printed using `accuracy_score()`.

Model Saving :

Model saved as `model.pkl`

Scaler saved as `scaler.pkl` for consistent preprocessing in deployment (like Flask apps).

Example Output:

python

Copy

Edit

Model accuracy: 86.89%

✓ Model and scaler saved.

✓ Why Logistic Regression is good for this:

Fast and interpretable.

Works well when the relationship between features and the outcome is linear.

Outputs probabilities (helpful in medical predictions).

Easy to deploy in real-world applications.