

Algorithm: Heart Disease Prediction using Machine Learning

Step 1: Data Collection

Import the dataset (e.g., UCI Heart Disease dataset).

Example: heart.csv

Step 2: Data Preprocessing

1. Load the data

```
pandas.read_csv()
```

2. Handle missing values

Fill or drop missing data.

3. Encode categorical variables

Convert categorical data to numeric using LabelEncoder or OneHotEncoder.

4. Feature Scaling

Normalize/Standardize features using StandardScaler or MinMaxScaler.

Step 3: Exploratory Data Analysis (EDA)

Plot histograms, heatmaps (correlation), and pair plots to understand relationships.

Identify important features for prediction.

Step 4: Splitting Dataset

Split dataset into training and testing sets:

```
train_test_split(X, y, test_size=0.2, random_state=42)
```

Step 5: Model Selection

Choose appropriate ML model:

Examples: Logistic Regression, Decision Tree, Random Forest, KNN, SVM

Train the model using:

```
model.fit(X_train, y_train)
```

Step 6: Model Evaluation

Predict using:

```
model.predict(X_test)
```

Evaluate with:

Accuracy, Precision, Recall, F1-score, ROC-AUC

Use `classification_report ()` and `confusion_matrix ()`

Step 7: Final Prediction System

Build a prediction function that takes input from the user and returns the prediction.

Save model using `joblib` or `pickle`.

Step 8: Deploy (Optional)

Build a simple web interface using:

HTML + Flask or Streamlit

Connect the UI to the trained model for the real time prediction.

FLOW:

Dataset->Preprocessing->Split->TrainModel->Evaluate->Tune->Predict->Deploy.