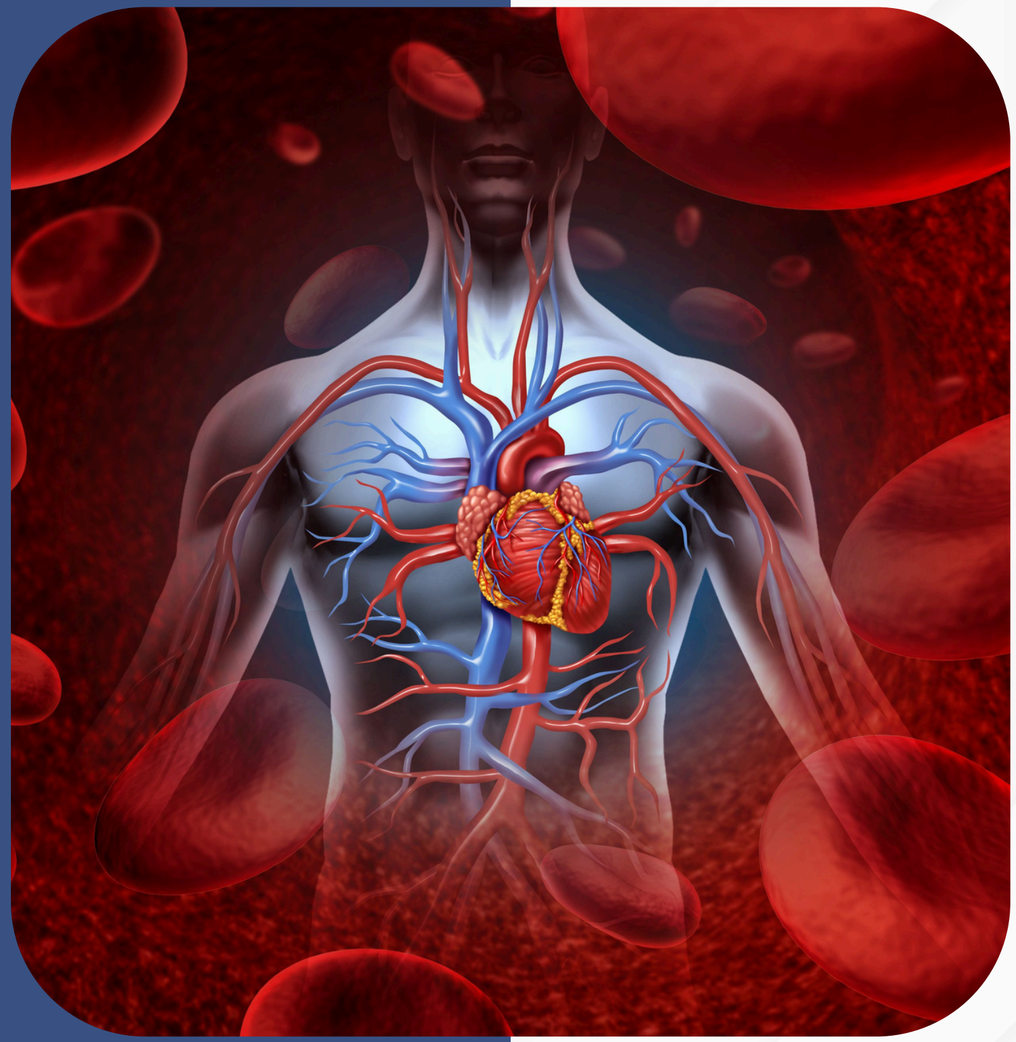
ARTIFICIAL INTELLIGENCE

CU6051NP

CVD PREDICTION

PRESENTED BY: SISIR PAUDEL
LONDONMETID : 23057059
DATE : 21/01/2026



Introduction

Cardiovascular Disease (CVD) is one of the most critical global health challenges. According to the World Health Organization, CVD accounts for nearly one-third of all global deaths. Artificial Intelligence enables early identification of high-risk patients, allowing preventive intervention and improved healthcare outcomes.

Topic Selection Justification

- Cardiovascular diseases are increasing globally, making early detection essential.
- AI offers effective solutions for prediction and diagnosis in healthcare.
- Availability of medical datasets enables practical model development.
- High real-world and academic relevance, ideal for impactful research.

AI Concepts Used

- Supervised machine learning for predictive modeling
- Classification algorithms to identify disease outcomes
- Data preprocessing & feature selection to improve model performance
- Performance evaluation using accuracy, precision, recall, and F1-score

Research Evidence

- Multiple studies show machine learning outperforms traditional statistical methods
- ML models provide higher accuracy in disease prediction
- Random Forest is effective due to its robustness and performance
- Logistic Regression is widely used for reliable predictions
- Both models offer good interpretability for medical use
- Proven suitability for cardiovascular risk assessment

Dataset Overview

- Patient demographic data: age and gender
- Clinical features: chest pain type and resting blood pressure
- Biochemical data: serum cholesterol and fasting blood sugar
- Diagnostic results: electrocardiogram (ECG) findings
- Exercise-related indicator: exercise-induced angina
- Target variable: presence or absence of cardiovascular disease

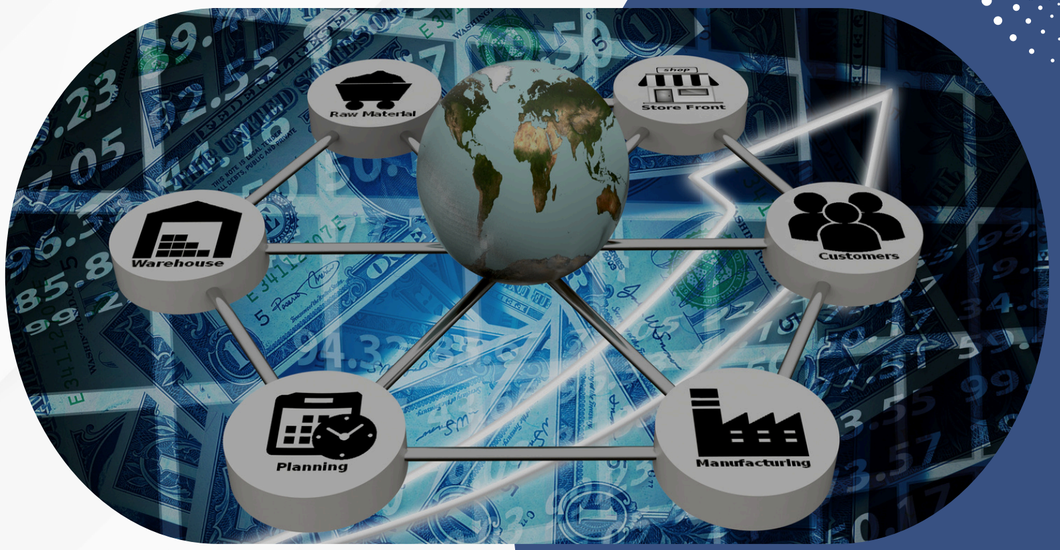
Data Preprocessing

- Handled missing values to ensure data quality
- Encoded categorical variables for model compatibility
- Scaled numerical features for uniformity
- Applied feature selection to remove irrelevant data
- Reduced dimensionality to simplify the dataset
- Improved model efficiency and performance

Solution Architecture

- Data collection from medical records
- Data preprocessing for quality and consistency
- Model training using machine learning algorithms
- Model evaluation with performance metrics
- Deployment for prediction in real-world use
- Ensures reliable and reproducible results

DevelopedApplication (How It Works)



Input Data

The system accepts patient health information, including demographic details and clinical test results, as input.

Data Processing

The trained machine learning model processes the input data, analyzing patterns and key risk factors related to cardiovascular health.

Risk Prediction

Using the analyzed data, the system predicts whether the patient is at risk of developing cardiovascular disease.

Clinical Support

The prediction results assist doctors in making informed decisions for diagnosis, prevention, and treatment planning.

Algorithms Implemented

Logistic Regression:

A widely used classification algorithm that models the relationship between input features and the probability of disease presence, serving as a strong baseline for performance comparison.

Decision Tree:

A rule-based machine learning model that splits data into branches based on feature values, allowing clear visualization and easy interpretation of decision-making processes.

Random Forest:

An ensemble learning algorithm that combines multiple decision trees to reduce overfitting, improve accuracy, and enhance overall prediction reliability.

Achieved Results

Logistic Regression:

- Accuracy: 0.96
- Precision: 0.96
- Recall: 0.96
- F1-Score: 0.96

Decision Tree:

- Accuracy: 0.96
- Precision: 0.96
- Recall: 0.96
- F1-Score: 0.96

Random Forest:

- Accuracy: 0.98
- Precision: 0.98
- Recall: 0.98
- F1-Score: 0.98

Real-World Impact

01

Early Diagnosis:

The AI-based system identifies early signs of cardiovascular disease by analyzing patient data, enabling doctors to take preventive or corrective actions at the right time.

02

Reduced Manual Analysis:

By automating data analysis and risk assessment, the system significantly reduces the need for manual evaluation by healthcare professionals, improving efficiency.

03

Improved Prediction Accuracy:

Advanced machine learning algorithms analyze complex patterns in medical data, resulting in more accurate and reliable predictions compared to traditional methods.

04

Enhanced Patient Outcomes:

Early and accurate predictions support better treatment planning, reduce complications, and contribute to improved overall patient health outcomes.

05

Hospital System Integration:

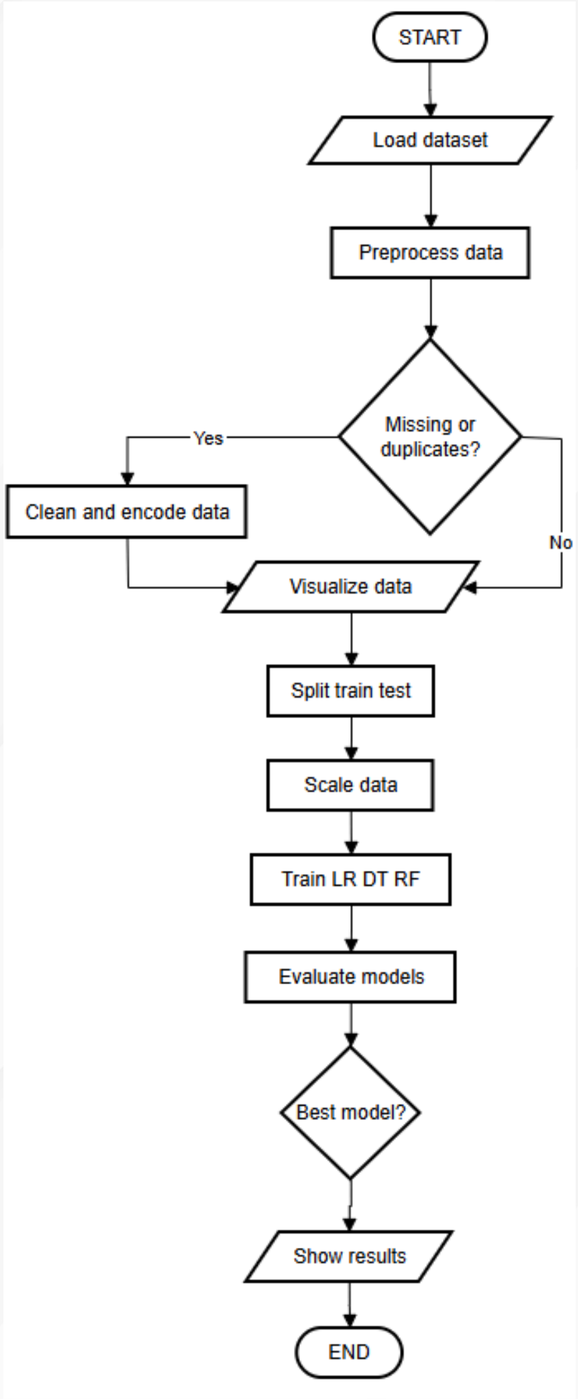
The system can be easily integrated into existing hospital information systems, functioning as a clinical decision support tool to assist doctors in real-time decision-making.

Pseudo Code Explaination

- Load Dataset: Import cardiovascular_disease.csv.
- Preprocess Data: Handle missing values, remove duplicates, and encode categorical features.
- Data Visualization: Plot feature distributions and correlation heatmap.
- Train–Test Split: Divide data into 80% training and 20% testing sets.
- Feature Scaling: Apply standard scaling to numerical features.
- Model Training: Train Logistic Regression, Decision Tree, and Random Forest models.
- Model Evaluation: Evaluate models using accuracy, precision, recall, and F1-score.
- Prediction: Generate predictions for healthy and high-risk inputs.
- Model Comparison: Compare performances and select the best model.

Flowchart

- Load and preprocess the dataset
- Handle missing values and encode features
- Visualize data patterns and relationships
- Split data into training and testing sets
- Scale numerical features
- Train and evaluate ML models
- Select the best-performing model
- Display final results



Conclusion

- **Effective Disease Prediction:** The project demonstrates that Artificial Intelligence can accurately predict cardiovascular disease using patient data.
- **Best Performing Model:** Random Forest emerged as the most reliable algorithm, offering high accuracy and robustness.
- **Enhanced Medical Decision-Making:** AI-driven healthcare systems can significantly support doctors by improving diagnostic accuracy and clinical decisions.

For Your Attention