HEART DISEASE PREDICTOR

Author: Dr.Mohanraj, SubhaSurya, Sudha, Sarath Kumar,

“Heart Disease Prediction using K Nearest Neighbour and K Means Clustering”, International Journal of Advanced Engineering Research and Science (IJAERS) 2016

Abstract:

Heart disease remains a leading cause of mortality worldwide. Early detection plays a critical role in reducing fatality rates and improving patient outcomes. This study develops a machine learning-based heart disease prediction model that leverages clinical data, such as age, blood pressure, cholesterol levels, and lifestyle habits, to assess the risk of heart disease. The proposed model integrates multiple algorithms, including logistic regression, decision trees, and neural networks, to enhance prediction accuracy. Results demonstrate the model's potential in assisting healthcare professionals with timely decision-making.

Introduction:

Heart disease is a significant public health challenge, accounting for millions of deaths annually. Despite advancements in medical diagnostics, late-stage detection often limits treatment effectiveness. With the rise of artificial intelligence and machine learning, predictive models have emerged as powerful tools in healthcare. These models analyze patient data to identify patterns and assess the likelihood of disease. This paper introduces a machine learning-based heart disease prediction system, discusses its clinical significance, and highlights the role of data-driven insights in early diagnosis and prevention.

Methodology:

1. **Data Collection**:
   Data was obtained from publicly available datasets, such as the UCI Heart Disease Dataset, encompassing parameters like age, gender, blood pressure, cholesterol levels, and ECG results.

2. **Preprocessing**:
   - Missing values were handled using imputation techniques.
   - Data normalization was performed to standardize feature scales.
   - Feature selection methods, such as correlation analysis, were used to identify the most relevant predictors.
3. **Model Development**:
   - Algorithms: Logistic regression, decision trees, random forests, and neural networks were implemented.
   - Training and Testing: The dataset was split into training (70%) and testing (30%) subsets.
   - Cross-validation was employed to ensure model robustness.
4. **Evaluation Metrics**:
   Metrics like accuracy, precision, recall, F1 score, and area under the ROC curve (AUC) were used to assess performance.

Conclusion:

The proposed heart disease prediction model achieved promising results, demonstrating its utility in clinical applications. By identifying at-risk individuals early, this tool can assist healthcare providers in tailoring interventions and reducing the burden of heart disease. Future work will focus on integrating real-time patient monitoring and expanding the dataset to improve generalizability across diverse populations.