BS in Software Engineering



Department of Software Engineering

Faculty of Computer Science & Information Technology

The Superior University, Lahore

FALL 2024

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Type (Nature of project) | | | [ ✓ ] **D**evelopment [ ] **R**esearch [ ] **R**&**D** | | |
| Area of specialization | | | Heart disease | | |
| **Project Group Members** | | | | | |
| Sr.# | Roll. # | Student Name | | Email ID | \*Signature |
| (i) | 101 | Muhammad Hassan | | Hassanafzaalas7@gmail.com | hassan |
| (ii) |  |  | |  |  |
| (iii) |  |  | |  |  |

\*The candidates confirm that the work submitted is their own and appropriate credit has been given where reference has been made to work of others

**Heart Disease Prediction**

**Introduction**

Heart disease is one of the leading causes of death worldwide. Early prediction of heart disease can significantly improve the prognosis and treatment strategies. This project aims to develop a predictive model using Logistic Regression to classify individuals as having heart disease or not based on their clinical features.

**Dataset Description**

The dataset used in this project contains information about individuals, including their clinical features and whether they have heart disease. Each row in the dataset represents one individual, and the columns represent features such as age, gender, cholesterol levels, and more.

Dataset File: heart\_disease\_data.csv

**Methodology**

The project follows these steps:

**1.** Load the dataset using Pandas.

**2**. Explore and preprocess the data, including checking for missing values.

**3.** Split the dataset into training and testing sets.

**4.** Train a Logistic Regression model using the training data.

**5**. Evaluate the model's performance using accuracy scores on training and test data.

**6.** Make predictions on new data instances.

**Implementation**

The Logistic Regression model was implemented using Python libraries such as Pandas, Numpy, and Scikit-learn. The model was trained on the dataset, and its performance was evaluated to ensure reliable predictions.

**Results**

The model achieved satisfactory accuracy on both training and testing datasets. It was able to predict whether an individual has heart disease or not based on input features.

**Prediction Example**

For a test case with the following features: Age: 62, Sex: 0, Chest Pain Type: 0, Resting Blood Pressure: 100, Cholesterol: 100, Fasting Blood Sugar: 0, Resting ECG: 0, Max Heart Rate: 160, Exercise-Induced Angina: 0, Oldpeak: 3.6, Slope: 0, Number of Major Vessels: 2, Thal: 2, the model predicted that the person has heart disease.

**Conclusion**

The Logistic Regression model provides a reliable tool for predicting heart disease based on clinical features. Future work can involve using more advanced models or larger datasets to further improve prediction accuracy.