D.Y. PATIL COLLEGE OF ENGINEERING & TECHNOLOGY,

KASABA BAWADA, KOLHAPUR

A

Project -I Synopsis

On

" Heart Disease Prediction System"

Presented by:

Sr No.	Name	Roll no.
1.	Bhagyashree Swami	17
2.	Rohit Chitari	09
3.	Aniket More	14
4.	Prathmesh Suryavanshi	15

Under the guidance of

Mr. N. M. Shinde

DEPARTMENT OF COMPUTER SCIENCE ENGINEERING

(ARTIFICIAL INTELLIGENCE & MACHINE LEARNING)

(2023-2024 TY-BTech. SEM V)

D.Y. PATIL COLLEGE OF ENGINEERING & TECHNOLOGY,

KASABA BAWADA, KOLHAPUR.

A

Project-I Synopsis

On

“ Heart Disease Prediction System”

Roll no.	Name	Signature
17	Bhagyashree Swami	
09	Rohit Chitari	
14	Aniket More	
15	Prathmesh Suryavanshi	

Date:

Place: Kolhapur

Mr. N. M. Shinde
(Project-I guide & coordinator)

Dr. S. V. Patil
(H.O.D. CSE AI-ML)

ABSTRACT

Cardiovascular diseases remain a significant global health concern, contributing to a substantial number of morbidity and mortality cases each year. Early and accurate prediction of heart disease can aid in timely interventions and personalized treatments. This study explores the application of machine learning techniques for heart disease prediction using a comprehensive dataset of clinical and diagnostic attributes.

The findings of this study indicate that machine learning algorithms can effectively predict heart disease based on the provided dataset. Certain algorithms demonstrate higher predictive accuracy and are particularly effective in identifying patients at risk. The study underscores the importance of feature selection in improving model performance and interpretability.

However, it is acknowledged that this research is limited by the nature and size of the dataset, and further validation using larger and diverse datasets is recommended. Additionally, the integration of expert medical knowledge and ethical considerations is crucial when deploying predictive models in real-world clinical settings. In conclusion, this study showcases the potential of machine learning in enhancing heart disease prediction accuracy. The results suggest that these techniques can complement traditional clinical assessment methods, enabling early intervention and personalized care for individuals at risk of cardiovascular diseases.

INDEX

CHAPTER 1. INTRODUCTION

Heart disease is one of the leading causes of death worldwide, and early detection plays a vital role in improving patient outcomes. With advancements in technology and the availability of large-scale medical data, predictive modeling has become a promising approach to identify individuals at risk of developing heart disease. In this context, a Heart Disease Prediction Program is designed to utilize machine learning algorithms and patient data to predict the likelihood of an individual developing heart disease in the future.

The primary objective of the Heart Disease Prediction Program is to assist healthcare professionals in making informed decisions and providing timely interventions to high -risk patients. By analyzing various risk factors and patterns from patient data, the program aims to generate accurate predictions, thereby enabling early intervention, personalized treatment plans, and lifestyle modifications.

The domain information of a heart disease prediction program refers to the specific aspects related to the field or area in which the program operates. In the case of a heart disease prediction program, the domain information includes:

Cardiovascular Medicine: The program operates within the domain of cardiovascular medicine, focusing on the prediction and prevention of heart disease. It considers various cardiovascular conditions, including coronary artery disease, heart failure, arrhythmias, and valvular heart disease.

Risk Factors: The program takes into account a range of risk factors associated with heart disease. These factors may include age, gender, family history, smoking habits, high blood pressure, high cholesterol levels, diabetes, obesity, physical inactivity, and stress.

Medical Data: The program utilizes medical data to predict the likelihood of heart disease. This data includes patient information such as medical history, lifestyle factors, symptoms, and results from medical tests and diagnostic procedures. Examples of relevant tests include blood pressure measurements, cholesterol profiles, ECG recordings, stress tests, and echocardiograms.

Machine Learning and Data Analysis: The program may incorporate machine learning algorithms and data analysis techniques to analyze the input data and identify patterns or correlations that are indicative of heart disease. These algorithms can be trained on large datasets of patient information to improve the accuracy of predictions.

Prevention and Treatment: The program aims to provide insights into the risk of developing heart disease, allowing healthcare providers to implement appropriate preventive measures and treatments. This may include lifestyle modifications (e.g., diet, exercise), medication management, and referral to specialists for further evaluation or intervention.

CHAPTER 2. LITERATURE REVIEW

[1] S. Ouyang, **"Research of Heart Disease Prediction Based on Machine Learning," 2022 5th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE), Wuhan, China, 2022, pp. 315-319, doi:10.1109/AEMCSE55572.2022**.

The use of massive clinical data in the medical field for supporting medical decision support is an inevitable development trend. Medical decision support is based on a variety of data sources accumulated and acquired in real-time in the clinic, and various machine learning algorithms are used to achieve classification of patient disease types or prediction of disease risks. This paper assists in performing cardiac disease prediction starting from different heart disease types (coronary heart disease) and data sets, summarizing the currently adopted machine learning diagnosis and prediction methods, highlighting the characteristics and differences of these methods, and analyzing the challenges and future developments. The results show that machine learning techniques have a wide range of applications in cardiac diseases. However, each machine learning method can only be applied to a specific scope due to the non-uniformity of medical data. At the end of the article, the prediction of heart disease is summarized.

[2] Soni, Jyoti, Ujma Ansari, Dipesh Sharma, and Sunita Soni. **"Predictive data mining for medical diagnosis: An overview of heart disease prediction."** ***International Journal of Computer Applications*** **17, no. 8 (2011): 43-48.**

Medical data mining has great potential for exploring the hidden patterns in the data sets of the medical domain. These patterns can be utilized for clinical diagnosis. However, the available raw medical data are widely distributed, heterogeneous in nature, and voluminous. These data need to be collected in an organized form. This collected data can be then integrated to form a hospital information system. Data mining technology provides a useroriented approach to novel and hidden patterns in the data.

[3] **"HEART DISEASE PREDICTION SYSTEM"** Mrs. Jayashree L K, Sushmita Makapur, T D Vyshnavi, T D Prathyusha, V K Kavya

We referred about data analysis from this research paper along with the data flow diagrams.

CHAPTER 3. PROBLEM STATEMENT

The high prevalence of heart disease and its significant impact on public health necessitate effective methods for early detection and risk assessment. Despite advancements in medical science, accurate and timely identification of individuals at risk of heart disease remains a challenge. This problem statement aims to develop a robust and reliable heart disease prediction model that leverages machine learning techniques to analyze comprehensive patient data and provide accurate risk assessment, facilitating proactive medical interventions and contributing to improved cardiovascular health outcomes.

The goal of this project is to develop a robust machine learning model for predicting the likelihood of heart disease in individuals based on a set of medical and lifestyle features. The model should analyze and interpret relevant data to provide accurate predictions and assist healthcare providers in making informed decisions about patient care and interventions. A comprehensive dataset containing a variety of attributes, including medical history, vital signs, blood tests, lifestyle choices, and possibly genetic information, will be provided. The dataset will be split into training and testing subsets to ensure unbiased model evaluation.

CHAPTER 4. OBJECTIVES

The objectives of the proposed work are as follows:

1. Process the Dataset
2. Train the model using Logistic regression Algorithm
3. Test the dataset
4. Check the accuracy
5. Compare with existing system
6. Deployment
7. To provide effective methods of Heart Disease Prediction to decision makers towards better heart health management.
8. To provide recommendation to Heart health management system.
9. To prevent heart disease.
10. To make people aware about a healthy life.
11. To predict the seriousness of disease.
12. To reduce the chances of upcoming heart disease.

CHAPTER 5. PROPOSED SYSTEM ARCHITECTURE

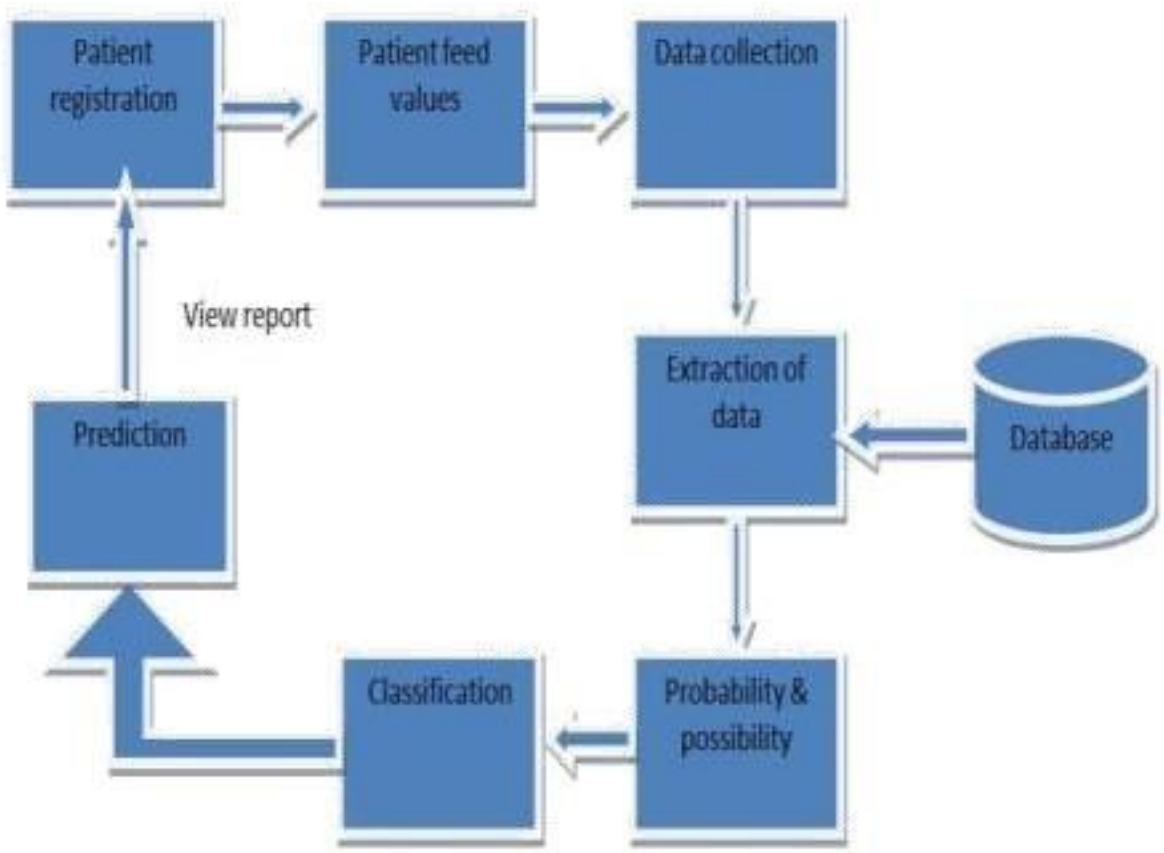

Fig 5.1. System Architecture

Initially the patient registers by providing certain parameters. That registered data is collected in a database by using machine learning techniques like data collection techniques and when he went to check about his health condition the collected values or data that has been stored in the database is been extracted by using some feature extraction techniques. When data is extracted, it under goes certain processes and therefore finally a disease is predicted and a report is generated. This is the overview of the heart disease prediction system using machine learning techniques.

Fig 5.2. Dataflow diagram.

This is the initial idea for the flow of the data. The data has to be flown from user to server and from server to the user for the prediction of the disease by entering details and sending the data. Communication is done between user and the server.

CHAPTER 6. MODULES

1. **Data Collection:** Gather a dataset containing relevant information about individuals, including their medical history, lifestyle factors, physiological measurements, and any diagnosed heart conditions. Ensure the dataset is diverse and representative of the population you aim to predict heart disease for.

2. **Data Preprocessing:** Clean the dataset by handling missing values, outliers, and inconsistencies. Perform feature engineering to extract meaningful features from the available data. This may involve transforming variables, normalizing data, or creating new features through domain knowledge.

3. **Data Split:** Divide the dataset into two or three subsets: a training set, a validation set, and optionally a test set. The training set is used to train the model, the validation set helps fine-tune the model's hyperparameters, and the test set provides an unbiased evaluation of the final model's performance.

4. **Model Selection:** Choose an appropriate machine learning algorithm for heart disease prediction. Some commonly used algorithms include logistic regression, decision trees, random forests, support vector machines (SVM), or neural networks. The choice of model depends on the dataset size, complexity, interpretability, and other specific requirements .

5. **Feature Selection:** Select the most informative features for heart disease prediction. This step can involve techniques like correlation analysis, recursive feature elimination, or regularization methods to identify the features that have the most significant impact on the prediction task.

6. **Model Training:** Train the chosen model using the training dataset. The model learns from the input features and their corresponding heart disease labels.

7. **Model Evaluation:** Assess the performance of the trained model using the validation set. Common evaluation metrics for classification tasks include accuracy, precision, recall, F1 score, and area under the receiver operating characteristic curve (AUC-ROC). Adjust the model's hyperparameters if necessary to optimize its performance.

8. **Model Testing:** Once satisfied with the model's performance, evaluate it on the independent test set. This provides an unbiased estimate of the model's generalization ability to predict heart disease on new, unseen data.

9. **Model Deployment:** If the model performs well on the test set, it can be deployed to make predictions on new data. The deployment can be as simple as providing a web interface or integrating the model into an existing healthcare system.

10. **Model Monitoring and Updating:** Continuously monitor the performance of the deployed model, and periodically update it to incorporate new data or account for changes in the population being predicted. This ensures that the model stays accurate and reliable over time.

These modules collectively create a heart disease prediction system that can help healthcare professionals identify individuals at risk of heart disease and provide timely interventions and treatments.

CHAPTER 7. SYSTEM REQUIREMENTS

1.HARDWARE REQUIREMENTS

- Processor: intel CORE i3
- 8 GB RAM
- 256 GB ROM

2.SOFTWARE REQUIREMENTS

- OS: Window OS
- Programming languages: python, HTML, CSS
- Database : Kaggle
- Internet

REFERENCES

[1] S. Ouyang**, "Research of Heart Disease Prediction Based on Machine Learning,"** 2022 5th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE), Wuhan, China, 2022, pp. 315-319, doi:10.1109/AEMCSE55572.2022.00071. https://ieeexplore.ieee.org/abstract/document/9948280/

[2] Soni, Jyoti, Ujma Ansari, Dipesh Sharma, and Sunita Soni. **"Predictive data mining for medical diagnosis: An overview of heart disease prediction."** *International Journal of Computer Applications* 17, no. 8 (2011): 43-48. https://www.academia.edu/download/79534142/5a18f6653b56138cd5196d20e2f39de189e3.pdf

[3] "HEART DISEASE PREDICTION SYSTEM" Mrs. Jayashree L K, Sushmita Makapur, T D Vyshnavi, T D Prathyusha, V K Kavya **http://www.irjcs.com/**

External Links:

https://towardsdatascience.com/

https://www.researchgate.net/

https://scholar.google.com/

Date:

Place: Kolhapur

Mr. Nitish Shinde (Project Guide & coordinator)

Dr. Siddheshwar Patil (H.O.D CSE AI-ML)