Counselling Code : VBIT
VIGNANA BHARATHI Institute of Technology
(A UGC Autonomous Institution, Approved by AICTE, Accredited by NBA & NAAC-A Grade, Affiliated to JNTUH)

HEART DISEASE PREDICTION USING MACHINE LEARNING

Dr. RAJU DARA
Professor
Dept., of Computer Science and Engineering

BATCH.NO : 23
B.GNANIKA - 21P61A0519
CH.SHIVANI - 21P61A0552
D.AMULYA - 21P61A0564

CONTENTS

ABSTRACT

- This project explores the development of a machine learning model to predict heart disease based on historical health data. By analyzing various health metrics, including blood pressure, cholesterol levels, and heart rate, the model aims to accurately forecast the likelihood of heart disease in individuals. The goal is to enhance early detection and enable timely intervention, potentially improving patient outcomes and reducing the risk of heart-related complications. Preliminary results demonstrate the model's capability to provide more precise predictions compared to traditional methods, emphasizing its potential in supporting preventive healthcare.Keywords: Heart Disease, Machine Learning, Prediction, Classification, Health Data, Early Detection, Preventive Healthcare.

INTRODUCTION

- Heart diseases
- Early detection and intervention
- Use of Machine learning algorithms in early prediction
- Supports healthcare professionals in diagnosing heart diseases early

LITERATURE SURVEY

Study	Approach	Pros	Cons
Study 1: ”Predicting Heart Disease Using SVM”	Support Vector Machine(SVM)	High accuracy in binary classification	Requires extensive tuning
Study 2: “Heart Disease Prediction with Neural Networks”	Neural Networks	Can capture complex patterns	Computationally intensive
Study 3: “Decision Trees for Heart Disease Prediction”	Decision Trees	Easy to interpret results	Prone to overfitting
Study 4: “ Logistic Regression for Heart Disease”	Logistic Regression	Simple and effective	May not handle complex relationships well

PROBLEM STATEMENT

- Issues in accurate prediction of heart diseases based on historical health data and various risk factors
- Limitations in terms of accuracy, interpretability, and generalization in existing systems

OBJECTIVE

- To develop a machine learning model to predict the likelihood of heart disease
- Identify the most effective method for prediction based on the performance
- Evaluate effectiveness and reliability of the model across different demographic groups

SYSTEM REQUIREMENTS

- **HARDWARE REQUIREMENTS:**
  - ***Processor:*** Intel CORE i5 or higher, Processor Speed
  - ***Memory:*** 16GB RAM minimum; 32GB recommended
  - ***Storage:***
    - Primary: 500GB SSD for OS and software
    - Data Storage: 1TB HDD or SSD, expandable

- **SOFTWARE REQUIREMENTS:**
  - ***Operating System:*** Windows 11
  - ***Database:***
    - ***Relational Database:*** My SQL8.0 or higher
    - ***Non-Relational Database(NoSQL):*** MongoDB for large-scale unstructured data
  - ***Programming Language:*** Python
  - ***Libraries:*** Scikit-learn, Pandas and NumPy, Matplotlib

PROPOSED METHODOLOGY

Data Collection

- Data Preprocessing
- Model Selection
- Training and Testing
- Evaluation Metrics

MODULES

- In this section, the results are presented and the accuracy outputs of the algorithms are displayed. A comparison is made between the algorithms based on their accuracy. When building predictive models in the domain of CVD prediction using ML algorithms, the training and testing data split ratio can vary based on a number of factors. In this work, the dataset sizes are 80 and 20 for training and testing respectively. The above results are obtained by using the in-built classifier present in the Sklearn – a machine learning library in python. By performing the train-test split and hyper parameter tuning, the model can be trained and tested on the given dataset. Additionally, this approach allows for obtaining the optimal parameter that yields higher efficiency. User interface can also be provided to obtain the values from the user to provide results directly to the interface used by the user.

IMPLEMENTATION

Data DescriptionDemographic

Data Pre-processing

Feature selection

Model Training

Machine learning Algorithms

Confusion Matrix

REFERENCES

1. https://www.kaggle.com/code/chayandatta/heart-disease-prediction-using-machine-learning
2. https://www.researchgate.net/publication/331589020_Heart_Disease_Prediction_System
3. https://www.academia.edu/52218642/Heart_Disease_Prediction_Using_Machine_Learning_Algorithms
4. https://www.scirp.org/journal/paperinformation?paperid=73781
5. https://thecleverprogrammer.com/2020/11/10/heart-disease-prediction-using-machine-learning/
6. https://thecleverprogrammer.com/2020/11/10/heart-disease-prediction-using-machine-learning/
7. https://www.sciencedirect.com/science/article/pii/S187705092031521O
8. https://towardsdatascience.com/predicting-presence-of-heart-diseases-using-machine-learning-36f00f3edb2c
9. Blake, C.L., Mertz : UCI Machine Learning Database-s http:/mlearn.ics.uci.edu/databases/ heartdisease/
10. UCI Machine Learning Repository. (2020). Heart Disease Dataset. Available at: https://archive.ics.uci.edu/ml/datasets/Heart+Disease

QUERIES ..?

THANK YOU..!