Project Proposal: Prediction of Heart Disease Using Machine Learning

Abstract

Heart disease is one of the leading causes of death worldwide, which makes early and accurate prediction crucial for effective intervention. In this project we will explore the use of machine learning (ML) techniques to predict heart disease based on clinical parameters. Predictive accuracy will be assessed using machine learning models, including Random Forest, Decision Trees, Support Vector Machines (SVM), and regression-based approaches. This study aims to compare model performance and determine the most effective approach for heart disease prediction. The dataset used for this research contains patient records with relevant health indicators.

Introduction

Cardiovascular diseases (CVDs) contribute significantly to global mortality rate, which makes accurate and timely diagnosis essential. Traditional diagnostic methods are often time-consuming and prone to errors (Shah et al., 2020). Recent advancements in ML techniques have demonstrated the potential to enhance predictive accuracy and assist in early disease detection (Alshraideh et al., 2024). This study aims to search which is the best ML-based predictive model for heart disease and assess its effectiveness compared to traditional methods.

Problem Statement

Early detection of heart disease remains a challenge due to the complexity of its risk factors and symptoms. Conventional diagnostic approaches depend heavily on clinical expertise and extensive testing. Machine learning provides an alternative method by leveraging large datasets to identify patterns and correlations that may not be apparent to medical professionals (Mall, 2024). This

project investigates the feasibility of using ML algorithms to predict heart disease with high accuracy.

Objectives

- To train and test predictive models using machine learning techniques for heart disease detection.
- To evaluate different ML algorithms, including Decision Trees, Random Forest, SVM, and Regression.
- To compare model performance based on key evaluation metrics such as accuracy, precision, recall, and F1-score.
- To identify the most influential features contributing to heart disease prediction.

Literature Review

There are several studies that have explored ML techniques for heart disease prediction. Ogunpola et al. (2024) developed ML-based predictive models for cardiovascular disease detection, demonstrating the effectiveness of deep learning techniques. Similarly, Yadav, Soni, and Khare (2023) assessed multiple ML models and found that ensemble methods yielded high accuracy. Ramalingam, Dandapath, and Raja (2018) conducted a survey of ML techniques used for heart disease prediction, highlighting the importance of feature selection and data preprocessing. Another study by Ali et al. (2021) compared supervised learning algorithms and concluded that ensemble models often outperform individual classifiers. These studies reinforce the potential of ML in healthcare applications.

Methodology

Dataset

The dataset used for this research consists of patient records containing medical attributes such as age, blood pressure, cholesterol levels, and other cardiovascular indicators. The data will be preprocessed to handle missing values and normalize numerical features.

Machine Learning Models

The following machine learning models will be used:

- **Decision Tree**: A simple and effective model for classification tasks.
- **Random Forest**: An ensemble technique that improves predictive performance by combining multiple decision trees.
- **Support Vector Machine (SVM)**: A model that is particularly useful for binary classification problems.
- **Regression:** An algorithm that predicts the probability of the output based on key features.

Model Evaluation

Each model will be evaluated using:

- **Accuracy**: Measures overall prediction correctness.
- **Precision and Recall**: Evaluate false positives and false negatives.
- **F1-Score**: Balances precision and recall.

Expected Outcome

This study anticipates that ML models will achieve high accuracy in predicting heart disease. The comparative analysis will highlight the most effective algorithm, providing insights into practical implementation in health care settings.

Conclusion

The proposed project seeks to leverage ML techniques for heart disease prediction, offering an efficient alternative to traditional diagnostic methods. By identifying the best-performing model, this study can contribute to the ongoing advancement of AI in medical diagnosis.

References

1. Ali, M. M., Paul, B. K., Ahmed, K., Bui, F. M., Quinn, J. M., & Moni, M. A. (2021). Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison. *Computers in Biology and Medicine, 136*, 104672. https://www.sciencedirect.com/science/article/abs/pii/S0010482521004662
2. Alshraideh, M., Alshraideh, N., Alshraideh, A., Alkayed, Y., Al Trabsheh, Y., & Alshraideh, B. (2024). Enhancing heart attack prediction with machine learning: A study at Jordan University Hospital. *Applied Computational Intelligence and Soft Computing, 2024*(1), 5080332. https://onlinelibrary.wiley.com/doi/full/10.1155/2024/5080332
3. Mall, S. (2024, May). Heart Attack Prediction using Machine Learning Techniques. In *2024 4th International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)* (pp. 1778-1783). IEEE. https://ieeexplore.ieee.org/abstract/document/10617300
4. Ogunpola, A., Saeed, F., Basurra, S., Albarrak, A. M., & Qasem, S. N. (2024). Machine learning-based predictive models for detection of cardiovascular diseases. *Diagnostics, 14*(2), 144. https://www.mdpi.com/2075-4418/14/2/144
5. Ramalingam, V. V., Dandapath, A., & Raja, M. K. (2018). Heart disease prediction using machine learning techniques: A survey. Retrieved from https://www.researchgate.net/publication/325116774_Heart_disease_prediction_using_machine_learning_techniques_A_survey
6. Shah, D., Patel, S., & Bharti, S. K. (2020). Heart disease prediction using machine learning techniques. *SN Computer Science, 1*(6), 345. https://iopscience.iop.org/article/10.1088/1757-899X/1022/1/012072/meta
7. Yadav, A. L., Soni, K., & Khare, S. (2023, July). Heart diseases prediction using machine learning. In *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)* (pp. 1-7). IEEE. https://ieeexplore.ieee.org/abstract/document/10306469