

# Heart Disease Prediction using Machine Learning

**Gaikwad Ahilya Ram and Mr. Ganesh Deshapande**

Gramin College of Engineering, Vishnupuri, Nanded, Maharashtra, India

**Abstract:** *This study explores the application of machine learning algorithms for heart disease prediction. Leveraging a dataset comprising various health parameters, we employ classification models to analyze and predict the likelihood of heart disease. Results demonstrate the effectiveness of the proposed approach, showcasing its potential in early detection and preventive healthcare measures.*

**Keywords:** Cardiovascular disease, Machine learning, Health data analysis

## I. INTRODUCTION

Heart disease remains a leading cause of mortality worldwide, emphasizing the critical need for advanced methods in early detection and prevention. In recent years, machine learning has emerged as a promising tool in the healthcare domain, offering the potential to enhance predictive analytics and risk assessment. This study delves into the application of machine learning algorithms for heart disease prediction, leveraging a diverse set of health parameters. By exploring this intersection of technology and healthcare, we aim to contribute to the development of efficient and accurate models for identifying individuals at risk, thereby facilitating timely interventions and improving overall cardiovascular health outcomes.

According to the world health organization, every year 1.2 million deaths occur worldwide due to heart disease. The load of cardiovascular disease is rapidly increasing all over the world from the past few years. A healthy lifestyle and easiest detection are only ways to prevent the heart related disease. Many researches have been conducted in attempt to pinpoint the most influential factors of heart disease as well as accurately predict the overall risk. The effective technologies in machine literacy that have established on specific training and testing with the support of python and python libraries. The system acquires training directly from data and skill, grounded on this training, testing should be done on as per needful algorithm. This project aims to predict heart disease by analyzing data of patients which classifies whether they have heart disease or not using machine learning algorithm

## II. OBJECTIVES AND SCOPE OF PROJECT

### Project Objectives

- The main objective of this project is to predict the heart disease of a patient using machine learning algorithms.
- Taking input parameters from the user, preprocessing the data and predicting whether the person is suffering from Heart disease or not.

### Project Scope

Here the scope of the project is that integration of clinical decision support with computer-based patient records could reduce medical errors, enhance patient safety, decrease unwanted practice variation, and improve patient outcome. This suggestion is promising as data modeling and analysis tools.

e.g., Machine Learning is an emerging approach that helps in prediction, diagnosis of a disease. Several machine learning (ML) algorithms have been increasingly utilized for heart disease prediction. We aim to assess and summarize the overall predictive ability of ML algorithms in heart diseases

## III. EXISTING SYSTEM

- Clinical decisions are often made based on doctor's intuition rather than on the knowledge rich data hidden in the database.
- This practice leads to unwanted biases, errors and excessive medical costs which affects the quality of service provided to patients.

- There are many ways that a medical misdiagnosis can present itself. Whether a doctor is at fault, or hospital staff, misdiagnosis of a serious illness can have very extreme and harmful effects
- The National Patient Safety Foundation Cities that 42% of medical patients feel they have had experienced a medical error or missed diagnosis

The existing systems for heart disease prediction using machine learning encompass various approaches and models. Commonly, researchers have employed traditional classification algorithms such as Decision Trees, Support Vector Machines, and Logistic Regression to analyzer datasets comprising patient health information. Some studies have explored ensemble methods like Random Forests to enhance prediction accuracy.

Furthermore, deep learning techniques, including neural networks, have gained traction for their ability to capture intricate patterns within complex datasets. Feature selection methods and dimensionality reduction techniques are often integrated to optimize model performance and interpretability.

Several existing systems also focus on integrating electronic health records (EHR) and diverse medical data sources, allowing for a comprehensive analysis of patients' health profiles. The utilization of big data analytics contributes to a more holistic understanding of risk factors and improves the overall reliability of predictive models.

Despite the progress, challenges persist, such as the need for larger and more diverse datasets, ensuring model interpretability in clinical settings, and addressing ethical considerations related to patient data privacy. The ongoing evolution of machine learning techniques and the collaborative efforts between the healthcare and data science communities continue to shape and refine existing systems for heart disease prediction.

#### **IV. PROPOSED SYSTEM**

- The main objective of this research is to develop a prototype intelligent Heart Disease Prediction System using Machine Learning technique. So, its providing effective treatments, it also helps to reduce treatment cost.
- K-Nearest Neighbour(K-NN) : In K-NN algorithm a data point is taken whose classification is not available, then the number of neighbours, k is defined. After that k neighbours are selected according to the lowest Euclidian distance between the selected data points and their neighbours. The selected data point is then classified into a category , which is same as the category which has majority of neighbours among the K neighbors.
- Random Forest: Random Forest works by constructing multiple decision trees of the training data. Each of the trees predicts a class as an output and the class, which is is the output of the greatest number of decision trees is taken as the final result, in case of classification. In this algorithm we need to define the number of trees we want to create.
- The dataset used in this experiment is taken from a website called Kaggle

#### **V. RESOURCES**

##### **Software Requirements:**

- Operating System: Windows 10.
- Coding Language: Python 3.
- Software used: Anaconda Navigator (Jupyter Notebook).

##### **Hardware Requirements:**

- Processor: Intel or high.
- RAM:4 GB(min).
- Space on disk: minimum 100mb.

##### **For running the application:**

- Device: Any device that can access the internet

## VI. CONCLUSION

The main objective of this research is to develop a heart prediction system.

The project involved analysis of the heart disease patient dataset with proper data processing and predicting result .

## VII. ACKNOWLEDGMENT

We express our gratitude to all those who contributed to the successful completion of the Heart Disease Prediction Using Machine Learning project. This endeavor would not have been possible without the support and collaboration of various individuals and institutions.

Our sincere thanks go to [List of Names or Institutions] for their invaluable guidance, expertise, and mentorship throughout the project. Their insights have significantly enriched our understanding of machine learning applications in healthcare and cardiovascular disease prediction.

We extend our appreciation to the participants who willingly provided their health data for this research, enabling us to build robust and reliable predictive models. Their contributions are instrumental in advancing our knowledge and fostering improvements in healthcare outcomes.

## REFERENCES

- [1]. Arul Prakash, V. Arul, A. Jagannathan, A Look at of Efficient and more Suitable Load Balancing Algorithms in Cloud Computing, IJERCSE, ISSN (Online) 2394-2320, Vol 5, Issue 4, April 2018 4.
- [2]. Brad Woodberg, Mohan Krishnamurthy et.al., Configuring Juniper Networks NetScreen & SSG Firewalls, Chapter 1 - Networking, Security, and the Firewall, <https://doi.org/10.1016/B978-159749118-1/50003-4> 5.
- [3]. <https://www.thewindowsclub.com/smb-port-what-isport-445-port-139-used-for> 6.Cheng, H., Yan, X., Han, J., & Hsu, C.-W. (2007).
- [4]. Discriminative Frequent Pattern Analysis for Effective Classification. 2007 IEEE 23rd International Conference on Data Engineering. doi:10.1109/icde.2007.367917