**HEART DISEASE PREDICTION USING DATA MINING TECHNIQUES**

**SUBMITTED BY:**

EMON SINGHA (20-42344-1)

ADIB AHMED (20-42333-1)

MUHAMMAD ABDULLAH (20-42322-1)

**SUPERVISED BY:**

DR. AFZORA NAHAR

**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

**FACULTY OF SCIENCE & TECHNOLOGY**

**AMERICAN INTERNATIONAL UNIVERSITY - BANGLADESH**

**2022**

# TABLE OF CONTENTS

[TABLE OF CONTENTS i](#_bookmark0)

[LIST OF ABBREVIATIONS ii](#_bookmark1)

1. [INTRODUCTION](#_bookmark2) 2
   1. [Problem Statement](#_bookmark5) 3
   2. [Research Questions](#_bookmark6) 3
   3. [Research Objectives](#_bookmark7) 3
2. [METHODOLOGY](#_bookmark10) 4

3. [WORK SCHEDULE](#_bookmark14) 5

[REFERENCES 5-6](#_bookmark15)

# LIST OF ABBREVIATIONS

# NN - Neural Network

# DT - Decision Tree

# NB - Naïve Bayes.

# INTRODUCTION

One of the most common diseases in the world today is cardiovascular disease. It is predicted that 17.9 million deaths, or 15% of all natural deaths, occurred as a result of it in 2017 [3]. By monitoring the levels of different health indicators such blood pressure, cholesterol, heart rate, and glucose level, such as the early stages of cardiovascular disease can be identified [3]. Cardiovascular disease has an impact on not only the health of individuals but also the economies and costs of the nation’s [4]. Several machine learning and data mining methods are now being created and investigated for the purpose of forecasting various sorts of diseases [2]. Similar to this, several data mining, machine learning, and hybrid algorithms are being researched, created, and examined in an effort to identify and forecast the early stages of cardiac disease [5–9]. The process of identifying or foreseeing heart illness from a patient's records is known as heart disease diagnosis. When a patient has multiple diseases, doctors may not be able to accurately diagnose them in a timely manner [10]. The authors of [11] examined a wide range of studies on the prognosis of cardiac illnesses from various years and came to the conclusion that data mining approaches are more accurate. The ability of classification algorithms to process extremely huge data sets makes them commonly used in the healthcare industry. Naive Bayesian, Support Vector Machine, Nearest Neighbor, Decision Tree, Fuzzy Logic, Fuzzy Based Neural Network, Artificial Neural Network, and Genetic Algorithms are some of the regularly utilized techniques in healthcare [1].

# 1.1 Problem Statement

The rate of heart diseases is increasing at an exponential rate. Along with this people today have a lack of exercise and are less active. For most of them recreation is just another movie in bed or anything technology based. These factors boosted the rate of heart diseases to an unfortunately high percentage.

# 1.2 Research Questions

1. Which method is more efficient?
2. Which method is easy to elaborate?
3. What are the recent advances in different in different machine learning approaches?

# 1.3 Research Objectives

1. To find the best way to find accurate result by best method
2. To detect the best method have used
3. To identify recent data mining approaches

# 2. METHODOLOGY

# Based on the results of our literature analysis, we came to the conclusion that the three methodologies listed below are more accurate and effective at classifying and predicting. Consequently, we tested out these three techniques:

1. Neural Network
2. Decision Tree
3. Naïve Bayes.
4. **Decision Tree**

The Decision Tree tested for a different number of testing data, how the accuracy can be improved by removing some of the attributes and testing again. After manipulating the dataset i.e. increasing and decreasing the training and testing data, we got the following results for the three data mining techniques for the prediction of heart disease.

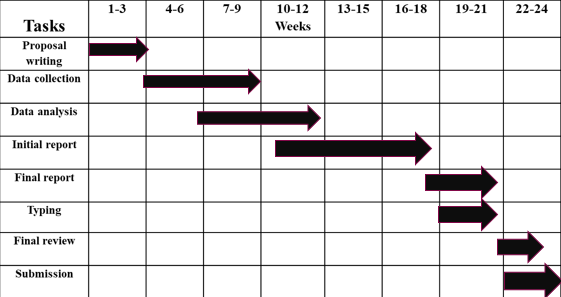
1. **Naïve Bayes**

The Naïve Bayes tested for different numbers of testing data, how the accuracy can be improved by removing some of the attributes and testing again.

1. **Neural Network**

The accuracy obtained for neural networks tested on different hidden layers, changing the number of epochs, increasing and decreasing learning rate and folds, and the activation functions. For improving the accuracy, we removed some of the attributes.

**3. WORK SCHEDULE**

  
  
  
  
  
**REFERENCE:**

1. Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective heart disease prediction using hybrid machine learning techniques. IEEE Access, 7, 81542-81554.
2. Tilve, A., Nayak, S., Vernekar, S., Turi, D., Shetgaonkar, P. R., & Aswale, S. (2020, February). Pneumonia Detection Using Deep Learning Approaches. In 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE) (pp. 1-8). IEEE.
3. Amin, M. S., Chiam, Y. K., & Varathan, K. D. (2019). Identification of significant features and data mining techniques in predicting heart disease. Telematics and Informatics, 36, 82-93.
4. Gokulnath, C. B., & Shantharajah, S. P. (2019). An optimized feature selection based on genetic approach and support vector machine for heart disease. Cluster Computing, 22(6), 14777-14787.
5. Das, Resul, Turkoglu, Ibrahim, et al.: Effective diagnosis of heartdisease through neural networks ensembles. J. Expert Syst. Appl.36, 7675–7680 (2009)
6. Das, Resul, Turkoglu, Ibrahim, et al.: Diagnosis of valvular heart disease through neural networks ensembles. J. Comput. Methods Progr. Biomed. 93, 185–191 (2009)
7. Gokulnath, C., Priyan, M. K., Balan, E. V., Prabha, K. R., Jeyanthi, R.: Preservation of privacy in data mining by using PCA based perturbation technique. In: Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), 2015 International Conference on (pp.202– 206). IEEE (2015)
8. Babaoglu, et al.: Assessment of exercise stress testing with arti- ficial neural network in determining coronary artery disease and predicting lesion localization. J. Expert Syst. Appl. 36, 2562–2566 (2009)
9. Rajeswari, K., et al.: Feature selection in ischemic heart disease identification using feed forward neural networks. Int. Symp. Robot. Intell. Sens. 41, 1818–1823 (2012)
10. Tarawneh, M., & Embarak, O. (2019, February). Hybrid approach for heart disease prediction using data mining techniques. In International Conference on Emerging Internetworking, Data & Web Technologies (pp. 447-454). Springer, Cham.
11. A. Powar, S. Shilvant, V. Pawar, V. Parab, P. Shetgaonkar and S. Aswale, "Data Mining & Artificial Intelligence Techniques for Prediction of Heart Disorders: A Survey," 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), Vellore, India, 2019, pp. 1-7, doi: 10.1109/ViTECoN.2019.8899547.

# 