

A Project Report
On
Heart Disease Prediction Based on
Self-Interested Test Reports
Submitted to
Acharya Nagarjuna University
In Partial Fulfilment of the Requirement for the Award of
Bachelor's Degree in
Information Technology
by

D. Naga Phanindra	Y19AIT428
G. Ramya	Y19AIT441
A. Lakshmi Prasanna	Y19AIT403
D. Beaula Rani	Y19AIT425

Under the guidance of
Mr. N. Srinivasa Rao
Asst. Prof, IT Dept.



Department of Information Technology
Bapatla Engineering College
(Autonomous)
Mahatmaji Puram, Bapatla - 522102
2022-2023
Affiliated to
Acharya Nagarjuna University

**A Project Report
on
Heart Disease Prediction Based on
Self-Interested Test Reports
Submitted to
Acharya Nagarjuna University**

**In Partial Fulfilment of the Requirement for the Award of
Bachelor's Degree in
Information Technology
by**

D. Naga Phanindra	Y19AIT428
G. Ramya	Y19AIT441
A. Lakshmi Prasanna	Y19AIT403
D. Beaula Rani	Y19AIT425

**Under the guidance of
Mr. N. Srinivasa Rao
Asst. Prof, IT Dept.**



**Department of Information Technology
Bapatla Engineering College
(Autonomous)**

**Mahatmaji Puram, Bapatla – 522102
2022-2023**

**Affiliated to
Acharya Nagarjuna University**

Bapatla Engineering College
(Autonomous)

Department of Information Technology

Mahatmaji Puram, Bapatla - 522102



CERTIFICATE

This is to certify that the Project entitled
**“Heart Disease Prediction Based on
Self-Interested Test Reports”**

D. Naga Phanindra	Y19AIT428
G. Ramya	Y19AIT441
A. Lakshmi Prasanna	Y19AIT403
D. Beaula Rani	Y19AIT425

is a record of bonafide work carried out by them, in the partial fulfillment of the requirement for the award of Degree of Bachelor of Technology (Information Technology) at BAPATLA ENGINEERING COLLEGE, Bapatla under the Acharya Nagarjuna University. This work is done during the year 2022-2023, under our guidance.

Date: / /

(Asst. Prof. Mr. N. Srinivasa Rao)

Project Guide

(Asst. Prof. Mr. M. Praveen Kumar)

Project Coordinator

(Dr. N. Sivaram Prasad)

H.O.D, I.T Department

Sig. of External Examiner

Acknowledgments

We are profoundly grateful to **Asst. Prof. Mr. N. Srinivasa Rao** for his expert guidance and continuous encouragement throughout to see that this project rights its target since its commencement to its completion.

We would like to express our deepest appreciation towards **Dr. Nazeer Shaik**, Principal, Bapatla Engineering College, **Dr. N. Sivaram Prasad**, Head of the Department of Information Technology and **Mr. M. Praveen Kumar**, Project Coordinator, whose invaluable guidance supported us in completing this project.

At last, we must express our sincere heartfelt gratitude to all the staff members of the Information Technology Department who helped us directly or indirectly during this course of work.

D. Naga Phanindra

G. Ramya

A. Lakshmi Prasanna

D. Beaula Rani

ABSTRACT

Heart-related diseases or cardiovascular diseases (CVDs) are the main reason for a huge number of deaths in the world over the last few decades and has emerged as the most life-threatening disease, not only in India but in the whole world. So, there is a need for a reliable, accurate, and feasible system to diagnose such diseases in time for proper treatment. Machine Learning algorithms and techniques have been applied to various medical datasets to automate the analysis of large and complex data. Many researchers, in recent times, have been using several machine learning techniques to help the healthcare industry and professionals in the diagnosis of heart-related diseases. Heart is the next major organ compared to the brain which has more priority in the Human body. It pumps the blood and supplies it to all organs of the whole body. Prediction of occurrences of heart diseases in the medical field is significant work. Data analytics is useful for prediction from more information, and it helps the medical center to predict various diseases. A huge amount of patient-related data is maintained monthly. The stored data can be useful for the source of predicting the occurrence of future diseases. Some of the data mining and machine learning techniques are used such as Artificial Neural Network (ANN), Random Forest, and Support Vector Machine (SVM). Prediction and diagnosing of heart disease become a challenging factor faced by doctors and hospitals both in India and abroad. To reduce the large scale of deaths from heart diseases, a quick and efficient detection technique is to be discovered. Data mining techniques and machine learning algorithms are very important in this area. The researchers are accelerating their research work to develop software with the help of machine learning algorithms which can help doctors to decide both prediction and diagnosing of heart disease. The main objective of this research project is to predict a patient's heart disease using machine learning algorithms.

Keywords: Neural Networks, Machine Learning, Supervised learning, Support vector machine, Random Forest.

Contents

1.	Introduction.....	1
	1.1 Project Scope	1
2.	Literature Survey.....	2
	2.1 Introduction to Literature Survey	2
	2.2 Existing System	5
3.	Requirement Specifications.....	6
	3.1 Software Requirement	6
	3.2 Hardware Requirement	6
	3.3 Libraries Requirement	6
4.	Requirement Analysis.....	8
	4.1 Machine Learning	8
	4.1.1 Supervised Learning	8
	4.1.2 Unsupervised Learning	8
	4.1.3 Reinforcement Learning	8
5.	System Design.....	9
	5.1 UML Diagrams	9
	5.1.1 Use Case Diagram	9
	5.1.2 Activity Diagram	11
6.	System Testing.....	12
	6.1 Types of Tests	12
	6.1.1 Unit Testing	12
	6.1.2 Integration Testing	12
	6.1.3 Functional Testing	12

6.1.4 System Testing	13
6.1.5 Module Testing	13
6.1.6 Acceptance Testing	13
7. Project Planning.....	14
7.1 Project Overview	14
7.2 Proposed System	14
8. Implementation.....	15
8.1 Problem statement	15
8.2 System Architecture	15
8.3 System Flow	16
8.3.1 Collection of datasets	16
8.3.2 Selection of attributes	16
8.3.3 Preprocessing of data	16
8.3.4 Prediction of disease	17
8.4 Dataset details	17
8.5 Algorithms	18
8.5.1 Support Vector Machine.....	18
8.5.2 Random Forest	18
8.5.3 TabNet	22
8.6 Implementation and Results	25
8.7 Source code	26
8.7.1 Input	37
8.7.2 Output	40
9. Screenshots of project.....	41

10. Conclusion and Future Scope.....	53
10.1 Conclusion	53
10.2 Future Scope	53
References.....	54

List Of Figures

5.1	Use Case Diagram.....	10
5.2	Activity Diagram.....	11
8.1	System Architecture.....	15
8.2	Preprocessing of data.....	16
8.3	Dataset.....	17
8.4	Samples of dataset.....	18
8.5	Support Vector Machine.....	21
8.6	Random Forest.....	21
8.7	Feature Block.....	22
8.8	Structure of TabNet algorithm.....	23
8.9	Feature Transformer Layer.....	23
8.10	Attentive Transformer layer.....	24
9.1	Types of Anginas.....	41
9.2	Cholesterol level.....	42
9.3	Values of the Slope variable.....	43
9.4	Dataset.....	44
9.5	First five records of dataset.....	44
9.6	Last five records of dataset.....	44
9.7	Attributes and their datatypes.....	45
9.8	Count of duplicates	45

9.9	Count plot of patients with (1) & without (0) heart disease..	46
9.10	Count plot of patients with fbs(1) & without fbs(0).....	46
9.11	Count plot for chest pain type	47
9.12	Count plot for Oldpeak.....	47
9.13	Scatter plot for svm classifier.....	48
9.14	Confusion matrix for SVM.....	49
9.15	Classification report for SVM.....	50
9.16	Confusion matrix for Random Forest.....	50
9.17	Classification report for Random Forest.....	51
9.18	Confusion matrix for TabNet.....	51
9.19	Classification report for TabNet.....	52