DPU

Dr. D. Y. Patil Unitech Society

DR. D. Y. PATIL INSTITUTE OF TECHNOLOGY

(Formerly Dr. D. Y. Patil Institute of Engineering and Technology)

Sant Tukaram Nagar, Pimpri, Pune.

Subject : Project Based Learning (PBL)

Branch : Artificial Intelligence And Data Science

Academic Year : 2021-22

Project Name : Disease Prediction System

# Synopsis

## Introduction:

Disease prediction using machine learning is a system where diseases are predicted by the machine learning algorithm on the basis of the symptoms provided by the user. In this project we will be collecting datasets, manipulate them and use various machine learning algorithms to create system that will efficiently predict the disease the user is suffering from on the basis of symptoms he/she possesses.

## Problem Statement:

Create a disease prediction system which will be used to predict heart diseases.

## Literature Survey :

1]

Name of the Paper :-IJERT-Heart Disease Prediction using Machine Learning

Authors:- Apurb Rajdhan, Avi Agarwal, Milan Sai, Dundigalla Ravi, Dr. Poonam Ghuli

Year:- 04-April-2020

Technique/Algorithms:-

Algorithms:-

1)Random Forest

2)Decision Tree

3)Logistic Regression

4)Naive Bayes

Performance Analysis:-

1)confusion matrix

Advantages:-

1)using random forest accuracy upto 90.6% can be achieved.

2)by testing with multiple algorithms we can get higher accuracy level for the same dataset

Disadvantages:-

1)Dataset is limited, higher accuracy can be achieved with more flexible inputs with a larger dataset.

## 2]

Name of the Paper :-Heart Disease Prediction using Machine Learning Techniques

Authors:- Devansh Shah, Samir Patel, Santosh Kumar Bharti

Year:- : 2 October 2020

Technique/Algorithms:-

Algorithms:-

1)Decision tree

2)Naïve Bayes

3)K-NN

4)Random forest

Performance Analysis:-

1)confusion matrix

Advantages:-

1)using random forest accuracy of 91% was achieved for people's dataset it was about 97%

2)by testing with multiple algorithms we can get higher accuracy level for the same dataset.

Disadvantages:-

1)Decision tree performed the worst by giving predictions with accuracy of only 42%.

2)Dataset is limited, higher accuracy can be achieved with more flexible inputs with a larger dataset.

3]

Name of the Paper :-Heart Disease Prediction using Hybrid machine Learning Model

Authors:- Dr. M. Kavitha , G. Gnaneswar , R. Dinesh , Y. Rohith Sai , R. Sai Suraj

Year:- : 26 February 2021

Technique/Algorithms:-

Algorithms:-

1)Decision tree

2)Random Forest Regression

3)Hybrid Model- Decision Tree + Random Forest combination

Performance Analysis:-

1)confusion matrix

Advantages:-

1)Using the hybrid of decision tree and random forest accuracy of 88% was achieved.

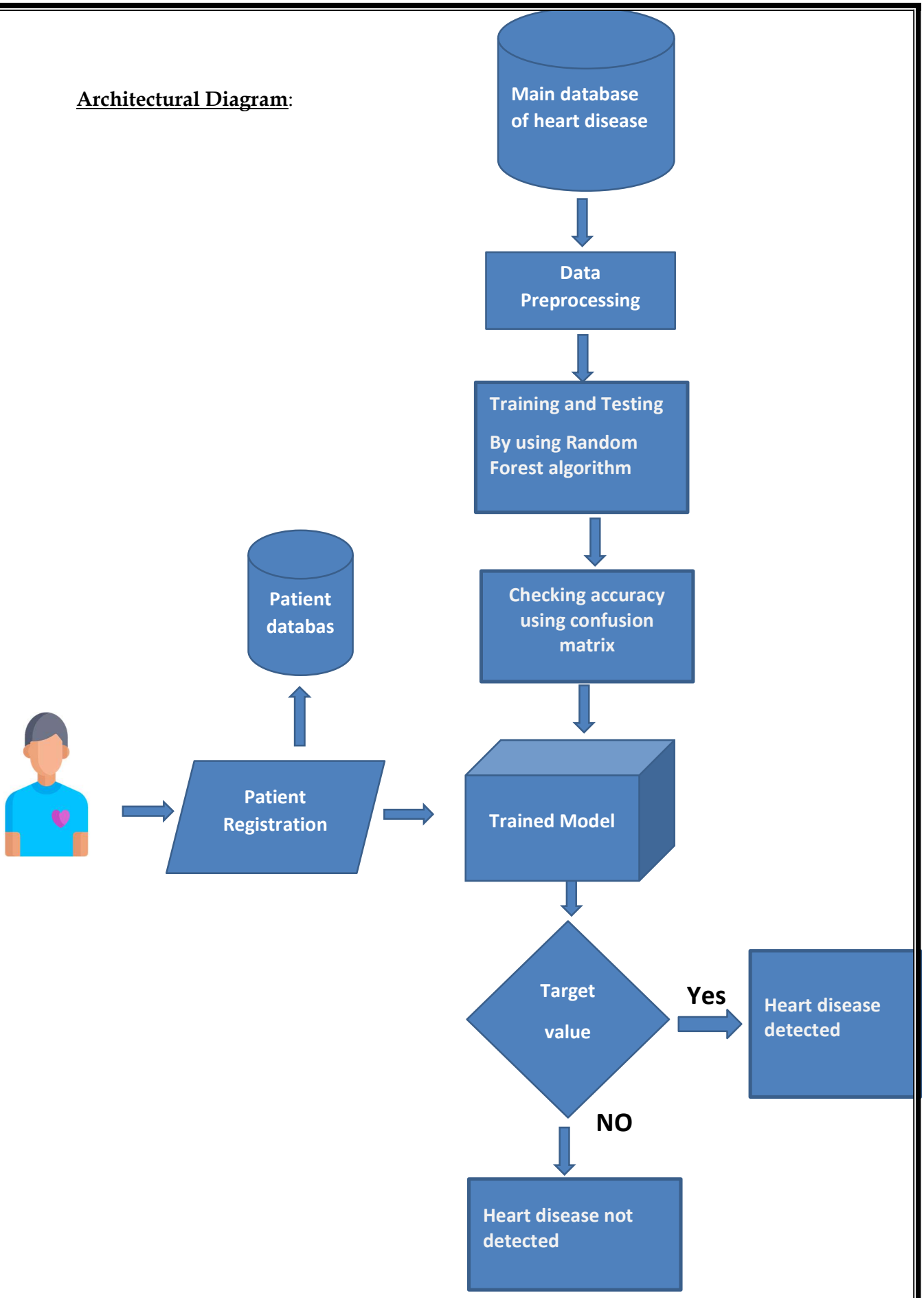
2)Due to use of combination of algorithms the ML model is more flexible to multiple different use cases.

Disadvantages:-

1) Hybrid algorithms of other classification algorithms are not used which may give better results.

2)Dataset is limited, higher accuracy can be achieved with more flexible inputs with a larger dataset.

**Architectural Diagram**:



## Future Scope:

In future we can add multiple models for predicting diseases and thus it can be use on a fairly wide scope for the analysis of the diseases. By which it will help to save many lives by predicting the disease at an early stage.

## Conclusion:

This system will help to predict the medical results efficiently. In this predicting system, we will provide a user-friendly interface that can be used by the users to detect whether their medical test results are positive or normal, i.e., it will detect the disease.

## References:


1. Google.com
2. Geeksforgeeks.com
3. Kaggle.com
4. Tutorialpoint.com
5. Javatpoint.com