ABSTRACT

Register No: 121006095 **Name:** Annapoorani.J

Heart disease is the most leading form of disease in today's generation. The lifestyle changes, improper food habits, lack of sleep, lack of exercise, addiction to drugs and alcohol and many more factors are the major causes of heart disease. Heart disease includes different kinds of disease such as Heart infections, Heart attack, strokes etc. There are many factors that cause heart disease, But I have considered the five most major factors of the cause. The major factors such as high blood pressure of blood, higher levels of cholesterol, improper levels of blood sugar, and a bad lifestyle are considered for the prediction of heart diseases. I designed the system with the help of a Jupyter notebook using fuzzy logic for prediction Heart disease diagnosis.

Technical Limitations & Ethical Challenges faced

Faced difficulties while designing the different kinds of rules that predicted the chances of getting the heart disease.

Keywords: fuzzy logic, heart diagnosis , prediction .

CHAPTER 1

INTRODUCTION

The heart is the major organ of the human body. Life of a human being completely depends on the functioning of the heart. Problem in heart reflects the major malfunctioning of organs such as the brain,liver, kidney and many more. Heart pumps the blood to the entire body, which carries the major fuel of the body called oxygen. Some of the heart diseases are shown in the figure [1.1].

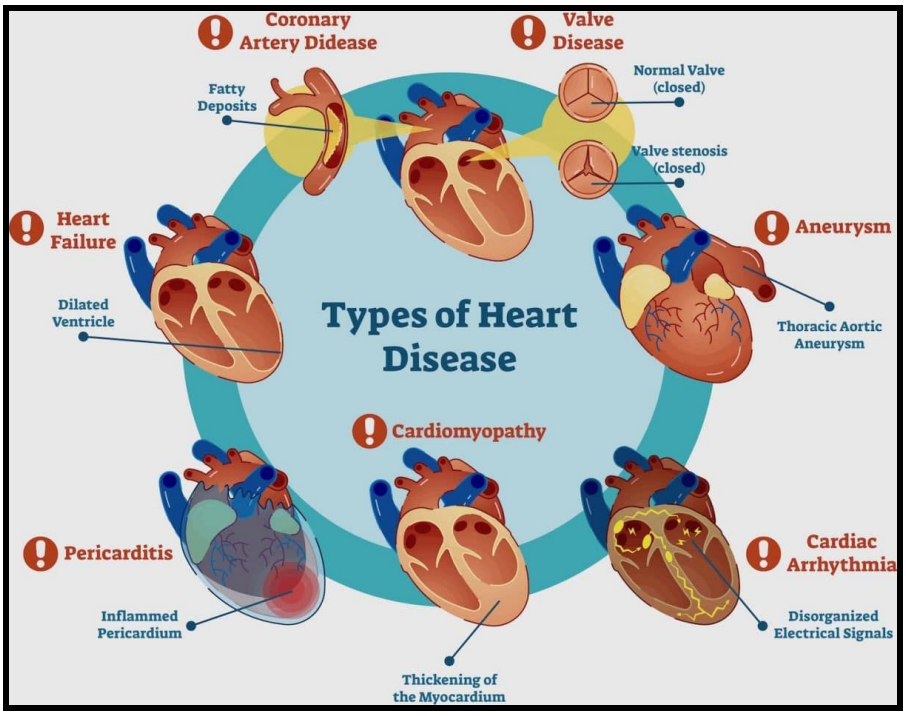


Figure 1.1: Types of heart disease

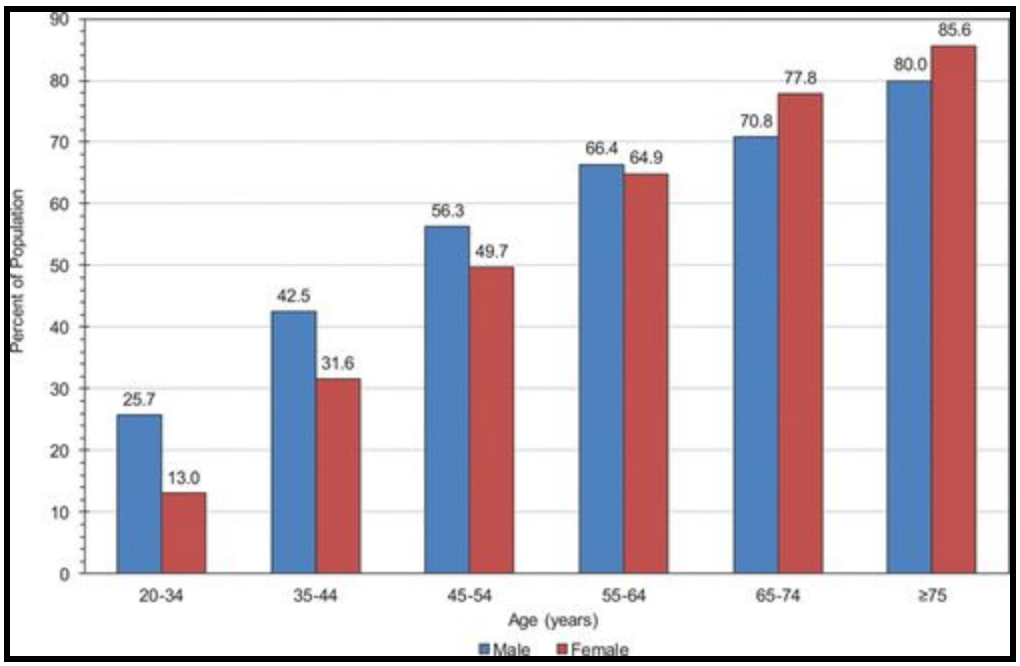
There are many causes of heart disease. some of the major causes of heart disease are listed below:

- Higher levels of blood pressure
- Family history
- Higher levels of cholesterol
- Improper sugar levels

- Physical inactivity
- Bacterial infection
- Other risk factors such as age,diet,gender etc.

Figure 1.2: causes of heart disease

The above figure[1.2] shows the causes of the heart diseases. The below figure[1.3] shows the statistics of heart disease in the year 2020. According to the data given by the World Health Organization about 17.9 million people die due to heart disease which is equivalent to 31% of deaths that occur worldwide in the year 2016.



Source: World Health Organization

Figure 1.3: Survey of heart disease in recent years

CHAPTER 2

LITERATURE OVERVIEW

Different types of methods are used for the prediction of heart disease. In [1] authors used a model which combines the Random Forest and multi-objective particle swarm optimization (MOPSO) which is a new approach for the prediction of heart disease in the patient. In [2] authors designed a smart method for prediction of heart disease using ensemble deep learning and the feature fusion approaches. Their model gives an accuracy of 98.5%. In [3] they introduced an enhanced performance level of prediction with an accuracy of upto 88.7% with the help of the model called hybrid random forest with a linear model (HRFLM). In [4] the authors designed a model which uses Advanced Encryption Algorithm for the prediction of the factors and risk of heart disease. In [5] authors an hybridization technique is introduced in which artificial neural network classifiers and decision tree are hybridized for enhanced performance of the prediction of heart disease.In [6] authors used two methodologies namely the Artificial neural network(ANN) by testing the input datasets and Case Based Reasoning(CBR) image similarity search by mapping the closeness of images of previous patients saved in database for the prediction of heart disease. In [7] authors did research on Data mining techniques such as K-nearest neighbor, Decision Tree, etc, used in the prediction of heart disease based on the input factors. They also considered factors such as thalach, food sugar level,number of vessel blockage, thallium scan and some more factors for their prediction. In [8] the authors used decision tree data mining method for the prediction of heart disease. In [9] the authors used a genetic algorithm with the help of back propagation approach for the prediction of the heart disease.In [10] authors delivered fuzzy logic using the matlab platform for the prediction of heart disease.

By studying all the works done by the authors which are stated above, I used a fuzzy logic method using Jupyter notebook for the prediction of heart disease.

CHAPTER 3

METHODS AND MATERIALS

fuzzy logic is often used in problems which are uncertain. It is best suitable for the systems having dynamic behaviour. Data sets are the main key for the fuzzy logic, using the below data we can predetermine the chances of getting heart disease or not. The system has five input attributes and one output attribute which gives the percentile chances of getting heart disease.

The steps to be followed in designing a fuzzy logic system is as follows.

1. Determining the input members and the output members.
2. Designing the member functions for every single input member.
3. Allotting the input members with the appropriate member function.
4. Allotting the output members with the member function.
5. Designing the required rules for the fuzzy system.

The five major factors that are considered in model are as follows:

- Blood pressure
    - Systolic
    - Diastolic
- High density lipoprotein(HDP)
- Low density lipoprotein(LDP)
- Triglycerides

Some of the other factors also include family history, physical inactivity, alcohol etc.

In the levels of blood pressure there are two subdivisions that are systolic and diastolic, In these there are three categories ,those are normal, risk and high. The ranges of the membership function is as follows,

Fuzzy set	Range Systolic	Range Diastolic

Normal	<120mmHg	<80mmHg
Medium	120-189mmHg	80-89mmHg
High	>140mmHg	>90mmHg

TABLE -I: .BLOOD PRESSURE LEVELS

Cholesterol is a main for the results and can alter the output with its small change in level. It consists of low density lipoprotein cholesterol, high density lipoprotein, and triglycerides. The fuzzy sets are as follows

Fuzzy set	Range
<40	Low
40-59	Normal
>40	Best

TABLE II.HIGH DENSITY LIPOPROTEIN

Fuzzy set	Range
<150	Low
150 - 499	Normal
>500	Best

TABLE - III: TRIGLYCERIDES RANGE

Fuzzy set	Range
<100	Normal
100-130	Close to normal
130-160	Close to high
160-190	Maximum

>190	Over Maximum

TABLE-IV: LOW DENSITY LIPOPROTEIN

CHAPTER 4

RULES FOLLOWED

The accuracy and quality of the fuzzy systems completely depends on the rules that are designed for the system. Here I have used 20 rules for designing my system. The rules are as follows

1. When LDP is very minimum and HDP is high ,triglycerides is minimum, systolic level is minimum and diastolic level is minimum – chances of getting heart disease is very low.
2. When LDP is minimum , HDP is high , triglycerides is minimum, systolic level is minimum and diastolic level is minimum –chances of getting heart disease is very low.
3. When LDP is nearly high , HDP is high , triglycerides is minimum , systolic level is minimum and diastolic level is minimum – chances of getting heart disease is medium.
4. When LDP is high , HDP is high , triglycerides is minimum, systolic level is minimum and diastolic level is minimum –chances of getting heart disease is medium.
5. When LDP is very high, HDP is high, triglycerides is minimum, systolic level is minimum and diastolic level is minimum – chances of getting heart disease is medium.
6. When LDP is very minimum, HDP is moderate, triglycerides is minimum , systolic level is minimum and diastolic level is minimum – chances of getting heart disease is medium.
7. When LDP is very minimum , HDP is minimum, triglycerides is minimum, systolic level is minimum and diastolic level is minimum – chances of getting heart disease is medium.
8. When LDP is very minimum, HDP is minimum, triglycerides is high, systolic level is minimum and diastolic level is minimum – heart disease is chance to occur.chances of getting heart disease is medium.
9. When LDP is very minimum, HDP is minimum, triglycerides is very high, systolic level is minimum and diastolic level is minimum –chances of getting heart disease is high.

10. When LDP is very minimum , HDP is minimum , triglycerides is very high, systolic level is average and diastolic level is minimum – chances of getting heart disease is high.
11. When LDP is very minimum, HDP is minimum, triglycerides is very high, systolic level is high and diastolic level is minimum –chances of getting heart disease is high.
12. When LDP is very minimum, HDP is minimum, triglycerides is very high, systolic level is minimum and diastolic level is minimum – chances of getting heart disease is high
13. When LDP is very minimum, HDP is minimum, triglycerides is very high, systolic level is very high and diastolic level is high – chances of getting heart disease is medium.
14. When LDP is very minimum, HDP is minimum, triglycerides is very high , systolic level is high and diastolic level is very high – chances of getting heart disease is low.
15. When LDP is minimum , HDP is high , triglycerides is minimum, systolic level is very high and diastolic level is minimum – chances of getting heart disease is low.
16. When LDP is very minimum, HDP is moderate, triglycerides is minimum , systolic level is minimum and diastolic level is minimum – chances of getting heart disease is low.
17. When LDP is very minimum, HDP is moderate, triglycerides is minimum, systolic level is minimum and diastolic level is minimum – chances of getting heart disease is medium.
18. When LDP is minimum, HDP is moderate, triglycerides is medium, systolic level is minimum and diastolic level is minimum – chances of getting heart disease is low.
19. When LDP is high , HDP is high, triglycerides is high and systolic level is minimum and diastolic level is minimum – chances of getting heart disease is high.
20. When LDP is high , HDP is high , triglycerides is high, systolic level is high and diastolic level is high – chances of getting heart disease is high.

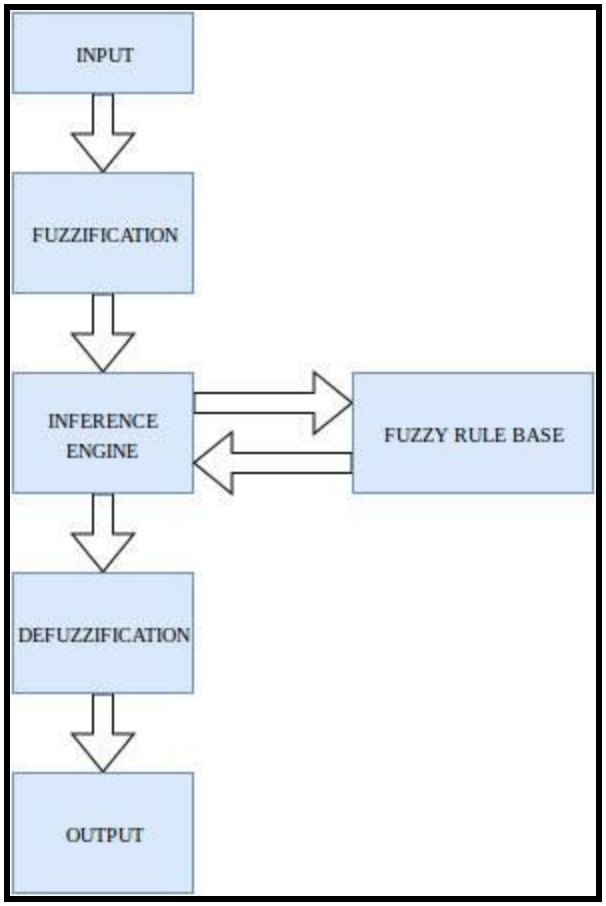


Figure-1.4 Fuzzy System Block diagram

CHAPTER 5

CODE AND DISCUSSION

The fuzzy logic code for heart disease prediction is as follows and it is carried out in jupyter notebook

In [14]:
```
import numpy as np
import skfuzzy as fuzz
from skfuzzy import control as ctrl
```

In [15]:
```
ldp = "low density lipid"
hdp = "high density lipid"
tri = "triglycerides"
sys ="systolic"
dia = "diastolic"
result = "result"
ideal="ideal"
results=0
```

In [16]:
```
ldp = ctrl.Antecedent(np.arange(0,190,1), ldp)
hdp = ctrl.Antecedent(np.arange(0,60,1), hdp)
tri = ctrl.Antecedent(np.arange(0,500,1), tri)
sys = ctrl.Antecedent(np.arange(0,140,1), sys)
dia = ctrl.Antecedent(np.arange(0,90,1), dia)
result = ctrl.Consequent(np.arange(0,100,1), result)
```

In [17]:
```
result.automf(3)
hdp.automf(3)
sys.automf(3)
dia.automf(3)
tri.automf(3)
ldp.automf(5)
```

Figure 1.5:code

In [18]:
```
rule1 = ctrl.Rule(ldp["poor"] & hdp["good"] & tri["poor"] & sys["poor"] & dia["poor"], result["poor"])
rule2 = ctrl.Rule(ldp["mediocre"] & hdp["good"] & tri["poor"] & sys["poor"] & dia["poor"], result["poor"])
rule3 = ctrl.Rule(ldp["average"] & hdp["good"] & tri["poor"] & sys["poor"] & dia["poor"], result["average"])
rule4 = ctrl.Rule(ldp["decent"] & hdp["good"] & tri["poor"] & sys["poor"] & dia["poor"], result["average"])
rule5 = ctrl.Rule(ldp["good"] & hdp["good"] & tri["poor"] & sys["poor"] & dia["poor"], result["average"])
rule6 = ctrl.Rule(ldp["poor"] & hdp["average"] & tri["poor"] & sys["poor"] & dia["poor"], result["average"])
rule7 = ctrl.Rule(ldp["poor"] & hdp["poor"] & tri["poor"] & sys["poor"] & dia["poor"], result["average"])
rule8 = ctrl.Rule(ldp["poor"] & hdp["poor"] & tri["poor"] & sys["poor"] & dia["poor"], result["average"])
rule9 = ctrl.Rule(ldp["poor"] & hdp["poor"] & tri["good"] & sys["poor"] & dia["poor"], result["good"])
rule10 = ctrl.Rule(ldp["poor"] & hdp["poor"] & tri["good"] & sys["average"] & dia["poor"], result["good"])
rule11 = ctrl.Rule(ldp["poor"] & hdp["poor"] & tri["good"] & sys["good"] & dia["poor"], result["good"])
rule12 = ctrl.Rule(ldp["poor"] & hdp["poor"] & tri["good"] & sys["good"] & dia["good"], result["good"])
rule13 = ctrl.Rule(ldp["mediocre"] & hdp["good"] & tri["poor"] & sys["good"] & dia["poor"], result["poor"])
rule14 = ctrl.Rule(ldp["mediocre"] & hdp["average"] & tri["poor"] & sys["poor"] & dia["poor"], result["poor"])
rule15 = ctrl.Rule(ldp["poor"] & hdp["average"] & tri["average"] & sys["poor"] & dia["poor"], result["poor"])
rule16 = ctrl.Rule(ldp["poor"] & hdp["good"] & tri["average"] & sys["poor"] & dia["poor"], result["poor"])
rule17 = ctrl.Rule(ldp["mediocre"] & hdp["average"] & tri["average"] & sys["poor"] & dia["poor"], result["average"])
rule18 = ctrl.Rule(ldp["poor"] & hdp["average"] & tri["average"] & sys["poor"] & dia["poor"], result["poor"])
rule19 = ctrl.Rule(ldp["good"] & hdp["good"] & tri["good"] & sys["poor"] & dia["poor"], result["good"])
rule20 = ctrl.Rule(ldp["good"] & hdp["good"] & tri["good"] & sys["good"] & dia["good"], result["good"])
```

In [19]:
```
rule_list=rule_list = [rule1, rule2, rule3, rule4, rule5, rule6, rule7, rule8, rule9, rule10, rule11, rule12, rule13, rule14, rul
result_ctrl = ctrl.ControlSystem(rule_list)
result_analysis = ctrl.ControlSystemSimulation(result_ctrl)
```

Figure 1.6:code

```
In [20]: result_analysis.input["low density lipid"] = 200
         result_analysis.input["high density lipid"] = 70
         result_analysis.input["triglycerides"] = 50
         result_analysis.input["systolic"] = 50
         result_analysis.input["diastolic"] = 2

In [21]: result_analysis.compute()

In [22]: results=(result_analysis.output["result"])

In [23]: print("probability of getting heart disease is:"+ str(results))
         print("RESULTS:")
         if(results<33.33333):
             print("Low chances of getting heart disease")
         if(results>33.33333 and results<66.6666):
             print("Chances of getting heart disease is moderate")
         if(results>66.6666):
             print("High chances of getting heart disease")

         probability of getting heart disease is:49.50000000000001
         RESULTS:
         Chances of getting heart disease is moderate
```

Figure 1.7:code

from the above code we can see one of the example where the LDP is 200, HDP is 70, triglyceride is 50 , systolic is 50 and diastolic is 2 which gives a percentile of getting heart disease as 49.5% which concludes a result of getting cancer is moderate. Likewise we can categorise as lower , moderate and higher chances of getting heart disease.

CHAPTER 6

CONCLUSION

From the above program we can see by using the data of Blood pressure ,cholesterol level and some more factors we can predict the occurrence of heart disease in three levels such as low, moderate and high. If the inputs values are maximum then the risk of getting heart disease is maximum. Other than fuzzy logic a different variety of neural network algorithms and methods can be implemented for the prediction of heart disease as mentioned in chapter 2. Having a healthy lifestyle, exercising regularly, avoiding stress, maintaining a proper diet, quitting smoking and alcohol are some of the ways we can avoid heart disease.