

HEART DISEASE DIAGNOSIS USING ARTIFICIAL NEURO FUZZY INFERENCE

MODEL(ANFIS)

BY

AIGBOKHAODE DESTINY OHIOLE

PSC2010366

**A PROJECT SUBMITTED TO THE DEPARTMENT OF COMPUTER SCIENCE,
FACULTY OF PHYSICAL SCIENCE, UNIVERSITY OF BENIN, BENIN CITY IN
PARTIAL FULFILMENT OF THE REQUIREMENT FOR AWARD OF THE BARCHELOR
OF T THE SCIENCE(B.Sc.) IN COMPUTER SCIENCE**

MAY, 2024

CERTIFICATION

This is to certify that this project was carried out by AIGBOKHAODE DESTINY OHIOLE with Matriculation Number: PSC2010366 of the department of Computer Science, University of Benin, Benin city.

DR. E.C IGODAN

(Project Supervisor)

Date

APPROVAL

This project is hereby approved in partial fulfillment of the requirements for the award of Bachelor of Science (B.Sc.) Degree in computer science at the university of Benin, Benin city.

Dr. E.C IGODAN

(Project Supervisor)

Date

PROF. G.O. EKUOBASE

(HEAD OF DEPARTMENT)

Date

DEDICATION

This work is dedicated to the Almighty God for his strength, wisdom, understanding, love, protection, inspiration and guidance. He granted me throughout my course of study.

ACKNOWLEDGEMENT

First and foremost, I want to express my gratitude to Almighty God for giving me with the strength and resources necessary to complete this project.

I would also like to acknowledge my supervisor, DR. E.C Igodan, who went out his busy schedule took out time to proofread and put a finishing touch to this work, deserves my heartfelt thanks for his guidance, patience, knowledge, and unwavering support during this project.

It is with glowing passion that I record my highest admiration and greatest gratitude to all of the lecturers of the Computer Science

Special thanks also to my friends/course mates, who in their own individual ways have imparted my life too with their words of knowledge all throughout my undergraduate days in the University. I am very much grateful.

I owe everything to my parents, Mr. and Mrs. Aigbokhaode; there is no doubt in my mind that they are the most important people in my life. I could not have gotten this far without their constant support and advice.

Finally, I would want to express my gratitude to every one of my classmates. Thank you all for all of your help during my time at university.

TABLE OF CONTENTS

TITLE PAGE	i
CERTIFICATION	ii
APPROVAL	iii
DEDICATION	iv
ACKNOWLEDGEMENT	v
ABSTRACT	viii
CHAPTER ONE	1
INTRODUCTION	1
1.1 Background of the Study	1
1.2 Research Motivation	3
1.3 Research Methodology	4
1.4 Research Aim and Objectives	5
1.5 Significance of the study	5
1.6 Scope of this Study	6
CHAPTER 2	7
LITERATURE REVIEW	7
2.1 OVERVIEW	7
2.2 INTRODUCTION TO CARDIOVASCULAR DISEASE	7
2.3 BAICS OF THE HEART	8
2.4 CARDIOVASCULAR DISEASES (CVDs)	8
2.5 CARDIOVASCULAR TREATMENTS	10
2.6 ARTIFICIAL INTELLIGENCE	11
CHAPTER THREE	21
SYSTEM ANALYSIS AND DESIGN	21
3.1 INTRODUCTION	21
3.2 SYSTEM ANALYSIS	21
3.3 SYSTEM DESIGN	21
3.4 OVERVIEW OF ANFIS AND GA	22
3.4.1 MATHEMATICAL FRAMEWORK OF ANFIS	22
3.5 TRAINING ANFIS BASED ON GENETIC ALGORITHM	20
3.6 CHROMOSOME CODING	21
3.7 FITNESS FUNCTION	22
3.8 SELECTION	23

3.8.1 CROSSOVER.....	23
3.9 DATASET AND FUZZY RULE GENERATION	24
3.10 FUZZY RULE GENERATION	25
3.12 EVALUATION METRICS	26
3.13 Results	26
3.14 Performance comparison	27
CHAPTER FOUR	28
SYSTEM IMPLEMENTATION AND DESIGN	28
4.1 INTRODUCTION	28
4.2 SYSTEM REQUIREMENTS	28
4.2.1 HARDWARE REQUAIREMENT:	28
4.2.2 SOFTWARE REQUIREMENT	28
4.3 MODEL DEVELOPMENT TOOL	29
4.3.1 PYTHON	29
4.3.1.4 ANACONDA NAVIGATOR	30
4.4 SYSTEM TESTING	31
4.6 TERMINAL EXECUTION	35
4.7 PYTHON CODE OUTPUT	36
CHAPTER FIVE	37
CONCLUSION, RECONMENDATION AND FUTURE WORKS	37
5.1 CONCLUSION	37
5.2 RECOMMENDATIONS	37
5.3 FUTURE WORKS SUMMARY	37
REFERENCE	39
4.5 PYTHON CODE	42

ABSTRACT

This project focuses on developing an intelligent system for heart disease prediction using Adaptive Neuro-Fuzzy Inference Systems (ANFIS) and Genetic Algorithm (GA). Cardiovascular disease is a leading cause of mortality worldwide, necessitating improved diagnostic tools. The proposed system leverages the strengths of neural networks and fuzzy logic to handle the complexities and uncertainties inherent in medical data. The ANFIS model is optimized using GA to enhance the accuracy of predictions. The system was tested on the UCI Cleveland heart disease dataset, achieving a satisfactory accuracy rate of 91.25%. This indicates the potential of the system to assist clinicians in making reliable diagnoses based on a combination of key medical attributes. Future work may involve extending the system to diagnose other complex diseases and incorporating additional heuristic algorithms to further refine model performance.

CHAPTER ONE

INTRODUCTION

1.1 Background of the Study

Cardiovascular diseases (CVDs) encompass a range of heart and blood vessel disorders. Types of coronary artery disease include; coronary artery disease (CAD), heart failure, arrhythmias, and heart valve problems. Coronary artery disease results from pathological changes in lining of arterial walls. These changes increasingly limit blood flow in coronary artery disease (ischemia) this deprives the myocardium of oxygen. Result from the same pathological that causes strokes, kidney damage and peripheral artery disease. This pathology is called Atherosclerosis. Symptoms associated with this typically are chest pain/ angina and other symptoms initially on exercise. (Ephzibah *et al.*, 2012).

As the condition develops, it can lead to a blockage which leads to myocardial infarction/ heart attack which results in death of myocardial tissue (sudden death in the individual) CAD is the single leading cause of death in Australia (12%) and most developed countries and also Survivors typically progress on to heart failure. There are three main coronary arteries: right coronary artery, circumflex artery, left anterior descending artery. (Ephzibah *et al.*, 2012).

Treatments of life-threatening CAD include: Surgical thrombectomy is a type of surgery to remove a blood clot from inside an artery or vein. Normally, blood flows freely through your blood vessels, arteries, and veins. Your arteries carry blood with oxygen and nutrients to your body.

Rapid restoration of blood flow relieves symptoms and minimizes myocardial damage.

Cardiovascular diseases like coronary artery disease (CAD) have become the leading cause of death globally, prompting the medical community to seek solutions for early identification, effective treatment, and prevention, with a focus on minimizing costs and also reduce the burden on healthcare systems. Early

detection of CAD is crucial as it often remains undetected until advanced stages due to the absence

of clear signs, advances in artificial intelligence, particularly through Artificial Neural Fuzzy Inference System (ANFIS) and Genetic Algorithms (GA), have shown promise in improving the diagnosis of heart diseases. (Ephzibah *et al.*, 2012).

ANFIS is an intelligent computing technique that combines the advantages of artificial neural networks and fuzzy logic principles to create a model that can simulate decision-making with uncertain and vague information. ANFIS uses a network structure that mimics the methods of fuzzy logic and can generate if-then rules to derive outputs from inputs. The system is adaptive, meaning it can adjust the rules and membership functions using learning algorithms from neural networks. Genetic Algorithm is a search and optimization technique based on the principles of natural evolution and genetics. It is commonly used to solve optimization and search problems by evolving solutions with regard to a fitness measure. GA operates using operators like selection, crossover, and mutation to evolve the solutions towards the best solution. The parameters within ANFIS, such as membership functions for the fuzzy logic system, can be optimized using GA, enabling the system to find the most effective solutions.

(Feng *et al.*, 2022).

These technologies leverage the strengths of fuzzy logic (handling uncertainty) and evolutionary algorithms (effective searching of large solution spaces), often leading to systems that can perform complex decision-making tasks effectively.

The study focuses on the effectiveness of ANFIS trained with genetic algorithms (GA) in diagnosing heart disease. The integration of ANFIS and GA aims to optimize diagnostic accuracy through advanced feature selection and rule-based classification. (Feng *et al.*, 2022)

Feature selection is the task of identifying and selecting a useful subset of pattern- representing features from a large set of features. The number of features increases as the dimensionality expands. The benefits of feature selection are facilitating data visualization and understanding,

reducing the measurement and storage requirements, reducing training time, defying the curse of dimensionality to improve prediction performance. The objective is finally to construct and select subsets of features that are useful to build a good predictor.

A subset of useful features will exclude redundant but retain relevant features. After generating the best feature subset, it is used for classification. Fuzzy set theory and fuzzy logic are highly suitable for developing knowledge-based systems in medicine for diagnosis of diseases. In this paper it is discussed about how genetic algorithms and fuzzy logic combine together for efficient and cost-effective diagnosis of heart disease.

1.2 Research Motivation

The motivation for creating a system to predict cardiovascular disease is driven by the pressing need to address the challenges posed by this leading cause of death worldwide. With the continuous increase in cardiovascular diseases, partly due to lifestyle, overwork, and dietary changes, traditional diagnostic methods like ECG often fail to detect abnormalities when no heart attack is occurring, making early identification of heart disease extremely difficult.

Fuzzy expert systems and other AI-based methods have shown promise in dealing with the complexity and vast amount of clinical data associated with diagnosing heart diseases. The application of Adaptive Neuro-Fuzzy Inference Systems (ANFIS), optimized with genetic algorithms, exemplifies a robust approach that simulates human reasoning under uncertain conditions, providing a highly accurate diagnosis. These systems offer strong fault tolerance and robustness, crucial for addressing nonlinear and uncertain environments typically found in medical diagnosis. (Ephzibah *et al.*, 2012).

Moreover, cardiovascular diseases have been persistently the top cause of mortality globally, emphasizing the importance of effective and early diagnosis systems. The growing body of research and technological advancements in artificial intelligence for medical applications supports the

development of such predictive systems, which can potentially save lives by allowing for timely and accurate predictions.

Creating a predictive system for cardiovascular disease is not only a technical challenge but also a critical societal need that addresses a significant public health issue, leveraging the capabilities of modern computational methods to enhance the predictive accuracy and manage the growing burden of cardiovascular diseases more effectively.

The motivation stems from the need to address the growing impact of cardiovascular diseases and the potential of machine learning methods, like ANFIS, to enhance diagnostic accuracy, reduce healthcare costs, and facilitate early detection and treatment, which are crucial in reducing morbidity and mortality associated with heart diseases.

The diagnosis of heart disease remains a critical area of medical research due to the complexity and severity of the disease. The need for effective diagnosis methods is evident, as heart disease continues to be a leading cause of mortality globally (feng *et al.*, 2022)

1.3 Research Methodology

The study utilized the Cleveland dataset from the UCI Machine Learning Repository, commonly used for creating and evaluating machine learning models for predicting heart disease. This dataset comprises medical records of 303 patients, each characterized by 13 input attributes and one output attribute that signals the presence of heart disease. Attributes include various medical details such as age, sex, type of chest pain, resting blood pressure, cholesterol levels, fasting blood sugar, among others. These comprehensive features aid in the development and training of models to precisely predict heart disease outcomes from patient medical profiles.

In this project study, various methods are involved in achieving its aim. Initially, the data was preprocessed to transform images into numerical formats appropriate for model analysis. An Adaptive Neuro-Fuzzy Inference System (ANFIS) model was then constructed and optimized using

techniques such as genetic algorithms and Genetic Algorithm optimizer. The final stage included a comparative analysis of the model's output, focusing on classification metrics like Recall, accuracy, and Area Under the Curve (AUC).

1.4 Research Aim and Objectives

The aim of this research is to develop an intelligent system for predicting heart disease using the Adaptive Neuro-Fuzzy Inference System (ANFIS) model, optimized by Genetic Algorithms. The specific objectives are as follows:

- a. Design an optimized GA-ANFIS model for the products of cardiovascular disease.
- b. Implement (a) above using python programming language.
- c. Use coronary artery disease data as case study.

1.5 Significance of the study

The prediction of heart disease through sophisticated models has significant benefits for both individuals and governments:

For Individuals:

- i. **Early Detection and Treatment:** Predictive models for heart disease can identify at-risk individuals early, potentially before they exhibit any clinical symptoms. This enables earlier interventions, which can significantly improve outcomes by managing the disease proactively rather than reactively.
- ii. **Personalized Healthcare:** With accurate predictions, individuals can receive more personalized healthcare plans based on their specific risk factors, leading to more effective management of their health.
- iii. **Increased Awareness and Self-Management:** Knowing one's risk level can motivate individuals to adopt healthier lifestyles and make informed decisions about their health, potentially leading to better overall health and reduced risk of heart disease.

For Government and Public Health:

- i. **Reduced Healthcare Costs:** Early diagnosis and prevention of heart disease can significantly reduce the cost burden on healthcare systems by lowering the incidence of expensive emergency interventions and long-term healthcare services.
- ii. **Improved Resource Allocation:** Predictive models help in better allocation of healthcare resources by identifying high-risk populations who may need more focused healthcare interventions, thus optimizing healthcare delivery and effectiveness.
- iii. **Data-Driven Policy Making:** The use of predictive analytics in healthcare can provide valuable insights into the effectiveness of policies and treatments, enabling governments to formulate data-driven health policies that are more likely to succeed in reducing the incidence of heart disease.
- iv. **Enhanced Treatment Outcomes:** With the ability to predict heart disease more accurately, hospitals can improve treatment outcomes through timely and precise interventions, reducing the rates of complications and readmissions

1.6 Scope of this Study

The study aims to advance the diagnostic methods for heart disease using an Adaptive Neuro-Fuzzy Inference System (ANFIS) optimized with a genetic algorithm. The research focuses on utilizing these sophisticated computational techniques to process and analyze CAD for predicting heart disease.

CHAPTER 2

LITERATURE REVIEW

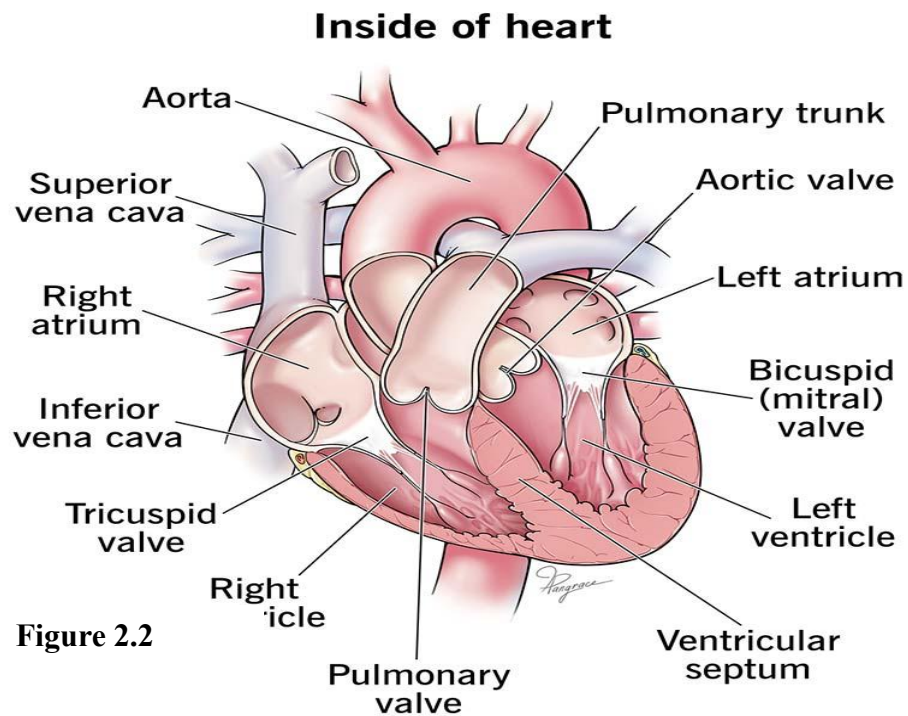
2.1 OVERVIEW

cardiovascular disease (CVD), highlighting its global prevalence and the necessity for effective diagnostics due to modern lifestyles. It explains the heart's function and details various CVD types, emphasizing the importance of early detection and management. The chapter also explores treatment options, including lifestyle changes, medications, and surgeries. A significant focus is on the role of artificial intelligence (AI) in improving diagnostic accuracy and healthcare outcomes, particularly through techniques like ANFIS and Genetic Algorithms. The chapter concludes with a summary table of studies, methods, and their contributions to cardiovascular health research.

2.2 INTRODUCTION TO CARDIOVASCULAR DISEASE

Cardiovascular disease remains a leading cause of mortality worldwide, with heart disease as its most prominent manifestation. According to the World Health Organization, heart disease has been the primary cause of death globally over the last two decades, emphasizing the crucial need for effective diagnostic methods. Modern lifestyles, which often include poor dietary habits and excessive stress, have led to a rise in the prevalence of cardiovascular diseases. This has spurred extensive research into AI-based diagnostics, where techniques like fuzzy logic have shown promise due to their robustness in handling uncertain and nonlinear problems. (Feng *et al.*, 2022).

2.3 BAICS OF THE HEART



The heart, a muscular organ located in the chest, functions as the circulatory system's pump, sending blood throughout the body to supply oxygen and nutrients while removing carbon dioxide and wastes. This blood flow supports all body functions and is crucial to maintaining homeostasis. (Feng *et al.*, 2022).

2.4 CARDIOVASCULAR DISEASES (CVDs)

They are primary cause of death worldwide, accounting for 32% of all global deaths in 2019, with an estimated 17.9 million fatalities. Heart attacks and strokes comprised 85% of these deaths. The majority of CVD-related deaths, over three quarters, occur in low- and middle-income countries. Additionally, CVDs were responsible for 38% of the 17 million premature deaths (under the age of 70) from noncommunicable diseases in that year. (Feng *et al.*, 2022)

Cardiovascular disease (CVD) refers to various conditions affecting the heart and blood vessels, often involving narrowed or blocked blood vessels that can lead to a heart attack,

chest pain (angina), or stroke. Other heart conditions, such as those affecting the heart's muscle, valves, or

rhythm, are also considered forms of heart disease.

Many cardiovascular diseases can be prevented by managing behavioral risk factors such as tobacco use, unhealthy diets, obesity, physical inactivity, and excessive alcohol consumption. Early detection and management of these risk factors are critical for preventing or delaying the onset of cardiovascular diseases. Treatment strategies can involve lifestyle changes, medication, and in some cases, surgery or other procedures. Understanding the heart's function and the factors contributing to cardiovascular health is vital in preventing disease and promoting overall well-being.

The following are types of cardiovascular diseases:

Coronary artery disease: A condition affecting the blood vessels that supply the heart muscle. **Cerebrovascular disease:** Involves the blood vessels that deliver blood to the brain.

Peripheral arterial disease: Affects the blood vessels that carry blood to the arms and legs.

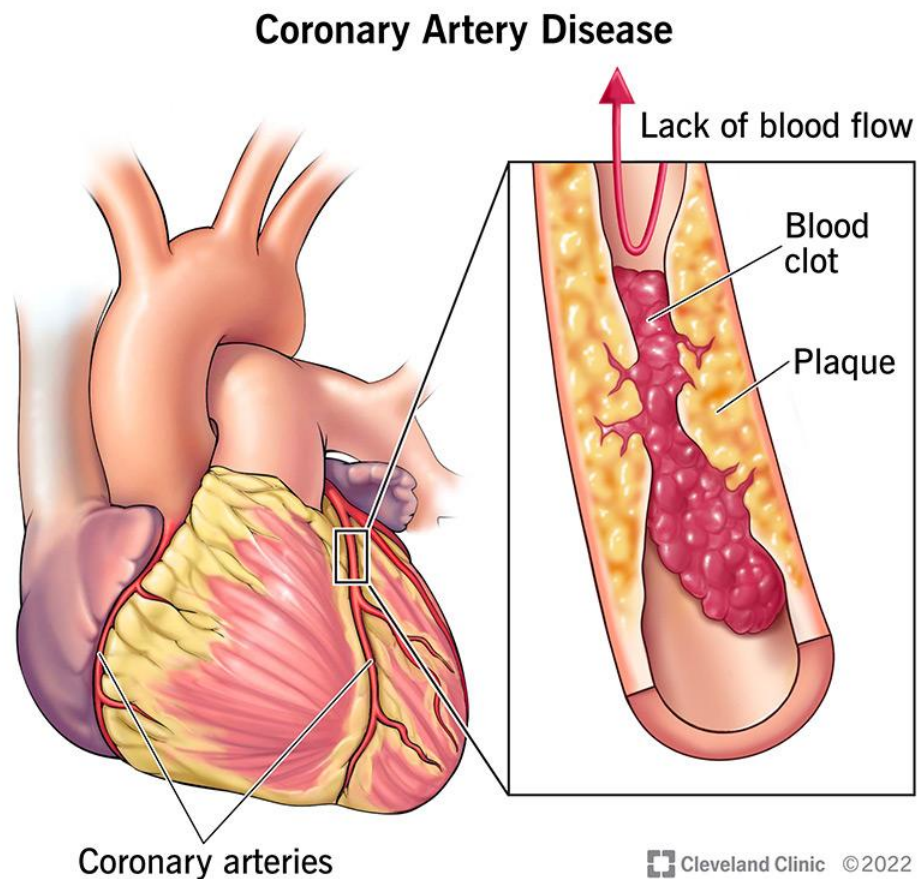
Rheumatic heart disease: Results from rheumatic fever and involves damage to the heart muscle and valves, triggered by streptococcal bacteria.

Congenital heart disease: Refers to heart defects present from birth, which alter the heart's normal development and function.

Deep vein thrombosis and pulmonary embolism: Conditions where blood clots form in the veins of the legs; these clots can break free and travel to the heart and lungs.

Heart attacks and strokes typically occur suddenly and are primarily caused by obstructions that stop blood flow to the heart or brain. The usual cause is the accumulation of fatty

deposits on the walls of blood vessels serving the heart or brain. Strokes may also stem from bleeding within brain blood vessels or clotting issues. Rajabzadeh Z. et al., (2023)



2.5 CARDIOVASCULAR TREATMENTS

The approach to treating cardiovascular disease can vary based on the specific symptoms and the type of disease present. Treatment options may include:

Lifestyle modifications: Such changes might involve dietary adjustments, increasing physical activity, and stopping the use of tobacco products, including e-cigarettes.

Medications: Depending on the specific cardiovascular condition, your healthcare provider might recommend certain medications to help control it.

Interventions or surgeries: Should medications be insufficient, various surgical interventions, such as placing stents in the heart or leg arteries, performing minimally invasive or open-heart surgeries, or conducting ablations or cardioversions, might be necessary.

Cardiac rehabilitation: This could involve a supervised exercise regimen designed to strengthen the heart.

Active surveillance: Sometimes, ongoing monitoring might be required, potentially without the need for medications or surgical interventions. Rajabzadeh Z. et al., (2023)

2.6 ARTIFICIAL INTELLIGENCE

Artificial Intelligence (AI) has made significant strides in advancing medical diagnostics and treatment, notably in the realm of cardiovascular health. The integration of AI, specifically through Adaptive Neuro-Fuzzy Inference Systems (ANFIS) and Genetic Algorithms (GA), has been pivotal in diagnosing conditions such as heart disease more accurately and efficiently. (Ephzibah *et al.*, 2012).

1. **Diagnosis and Prediction Accuracy:** AI systems, particularly those using ANFIS integrated with GA, have demonstrated high accuracy in diagnosing heart attacks. These systems utilize complex datasets to predict cardiovascular risks with a high degree of precision, as seen in multiple studies where the diagnostic accuracy reached as high as 98%.
2. **Explainability and Decision Support:** One of the critical features of modern AI systems in healthcare is their ability to provide explainable results. Systems designed for diagnosing heart disease often include explainable interfaces that allow clinicians to understand the diagnostic decisions made by the AI. This feature is crucial for clinical acceptance, where decisions need to be transparent and justifiable.

3. Enhanced Healthcare Outcomes: By predicting cardiovascular risks accurately, AI systems can potentially lead to better patient management and treatment outcomes. The integration of AI in medical diagnostics helps in early detection and timely intervention, which are critical for diseases like heart attacks where early treatment can significantly alter outcomes.

4. Innovation in Medical Procedures: AI applications in healthcare are not just limited to diagnostics but also extend to treatment and patient management. AI-driven models can simulate various treatment scenarios and predict outcomes, thus helping healthcare providers make informed decisions about the best treatment protocols.

5. Future Prospects: The ongoing research and development in AI for healthcare promise even more innovative solutions in the future. With advancements in machine learning and data analysis, the potential for AI to further enhance diagnostic accuracy and treatment efficacy continues to grow.

TABLE 2.1**SUMMARY OF REVIEWED LITERATURE**

AUTHOR	METHOD	DATA SOURCE	CONTRIBUTION TO KNOWLEDGE	LIMITATIONS
Rajabzadeh Z. <i>et al.</i> (2023)	ANFISF, FDA, LR.	MUMS	83.4% Accuracy (ANFIS) 72.4% Accuracy (LR) 77.7% Accuracy (FDA)	Imbalanced datasets posed a difficult classification problem, because they decreased classification performance.
Rajinikanth N. & Dr. Pavithira L. (2019)	ANFIS	UCI	92.5% Accuracy ANFIS	Little sample was used from the data set.
Rezaei-Hachesu P. <i>et al.</i> (2020)	FMDSS	UCI	94% Accuracy (FMDSS)	Time is a major obstacle.

Taylan O. <i>et al.</i> (2023)	SVR, MAR, ANFIS	Family medicine and cardiology clinics at a university hospital in Saudi Arabia.	94.7% Accuracy (ANFIS) 88.0% Accuracy (SVM)	<ol style="list-style-type: none"> 1. Clinical data are typically unbalanced and heterogeneous, and the data are arbitrary. 2. A high degree of statistical variation. 3. Arbitrary and varied processes and methods used for selecting criteria. 4. Few samples were utilized.
Bhagat R. K. (2022)	FL, ANN	UCI	89% Accuracy (FL)	Uneasiness in getting real time data.

Karsidani S. D. <i>et al.</i> (2022)	ANFIS, PSOM, PCI	EMC	98% Accuracy (ANFIS)	PCI prediction models have focused mainly on in-hospital or short-term outcomes
Abidin B. <i>et al.</i> (2019)	FNN, LR	Randomly chosen people at local health clinics, in the state of Kelantan, Malaysia.	79.8% Accuracy (FNN) 79.2% Accuracy (LR)	It has not been tested on real patients.
Sreedran S.. <i>et al.</i> (2023)	FIS	MATLAB	82.2% Accuracy. (FS)	Struggle to obtain up-to-date data.
MD. Ali L. <i>et al.</i> (2024)	MIS	UCI	98.08% Accuracy (MIS)	It cannot be enhanced further due to insufficient

				number of real patients in the field.
Rufai A. <i>et al.</i> (2018)	ANN, MLP.	UCI	92.2% Accuracy achieved	Hasn't being tested on real time patients.
Marateb H.R & Goudarzi S. (2015)	MLR, SFS	UCI	84% Accuracy (MLR)	Limited performance enhancement.
Ziasabounchi N. & Askerzade I. (2014)	ANFIS, HLA	UCI	92.30% Accuracy (ANFIS)	Due to less input parameters for different databases, the simulation cannot be further developed than this for now.
Warhade S.R (2017)	ANFIS	UCI	97.61% Accuracy	Not tested on real time

				patients.
Abushariah M.A.M <i>et al.</i> (2014)	MLP, ANFIS.	UCI	87.04% Accuracy (ANFIS) 75.93% Accuracy (MLP)	Obstacle encountered when trying to procure real- time data.
E.P Ephzibah <i>et al.</i> (2012)	FS, FES, GA	UCI	90% Accuracy (FES)	Not tested on real time patients
Subbulakshmi S. <i>et al.</i> (2019)	S-ANFIS	Not Given	92% Accuracy (S- ANFIS)	It is limited to two criteria only namely; Systolic Blood Pressure and Total Cholesterol.
Aghamohammadi M. <i>et al.</i> (2019)	GA, ANFIS	UCI	84.4358% Accuracy (ANFIS) 80% Accuracy (GA)	Missing features in large data sets.

Keikhosrokiani P. <i>et al.</i> (2022)	ANFIS by ABC	PhysioNet- Challenge	93% Accuracy (ABC-ANFIS) 85% Accuracy (ANFIS)	Obstacle encountered when attempting to procure real- time data
Adeli A. & Neshat M. (2010)	MIS	UCI	94% Accuracy (MIS)	Unavailability of real-time data.
Feng J. <i>et al.</i> (2010)	ANFIS and GA	UCI	91.25% Accuracy (ANFIS)	Generation of rule base will be very large.

LEGENDS

ANFIS - Adaptive Neuro Fuzzy Inference System

FDA - Flexible Discriminant Analysis

LR - Logistic Regression

MUMS - Mashhad University of Medical Sciences

UCI – University of California Irvine

HLA - Hybrid Learning Algorithm

SVR - Support Vector Regression

MAR - Multivariate Adaptive Regression

SFS - Sequential Fuzzy System.

GA - Genetic Algorithm

ANN – Artificial Neural Network

ABC - Artificial Bee Colony

FIS – Fuzzy Inference System

FS – Fuzzy System

EMC - Ekbatan Medical Center

FS - Feature selection

FES - Fuzzy Expert System.

PCI - Percutaneous Coronary Intervention

PSOM - Particle Swarm Optimization model.

FMDSS - Fuzzy Mobile Decision Support System.

S-ANFIS - Sugeno-Adaptive Neuro Fuzzy Inference

ANN - Artificial Neural Network

MIS - Mamdani inference system

CVD – cardiovascular disease

PSOM - Particle Swarm Optimization model.

MLP - Multiple layer perception.

MLR - Multiple logistic Regression

CAD - coronary artery diseases

MATLAB – Matrix Laboratory

CHAPTER THREE

SYSTEM ANALYSIS AND DESIGN

3.1 INTRODUCTION

System analysis involves gaining insights through determining system requirements and creating a logical design. Analyzing the current system identifies gaps, which are addressed in the design of a new system. The importance of system analysis and design can be summarized as follows:

- a) It provides exposure to various methodologies, tools, and techniques for developing new systems.
- b) It enhances analytical and problem-solving abilities.
- c) It involves learning about fact-finding and data collection methods.
- d) It is essential for the successful development and deployment of better systems.

3.2 SYSTEM ANALYSIS

The term "analysis" is used because it involves decomposing a whole into its components to better understand the characteristics, functions, and interrelations of these parts. The analysis process typically consists of three stages:

- i. Comprehending the current situation.
- ii. Identifying potential improvements.
- iii. Specifying the requirements for the new system.

3.3 SYSTEM DESIGN

System design is the determination of the overall system architecture consisting of a set of physical processing components, hardware, software, people, and the communication among them that will satisfy the system's essential requirements. It is a process of planning a new system or replacing an existing system by defining its components and modules to satisfy the specific requirements. The

proposed system seeks to compare ANFIS and GA methods to suggest which is best for detection and classification of Cardiovascular disease

3.4 OVERVIEW OF ANFIS AND GA

3.4.1 MATHEMATICAL FRAMEWORK OF ANFIS

ANFIS is a suitable intelligent computing technique based on Takagi-Sugeno fuzzy model. Combining the advantages of artificial neural networks and fuzzy logic, ANFIS not only acquires learning and adaptive ability, but also ensures that the weight values obtained can be explainable. The architecture of ANFIS consist of five layers. To better understand the system, a basic ANFIS structure consisting of two inputs and one output is shown in Figure2.2.1. By analysing Figure.1, it is shown that this structure can generate up to four rules that illustrated as follows:

Rule n: IF x is A_i and y is B_j , THEN $f_n = p_n + q_n + r_n \quad i, j \in \{1, 2\}$ and $n \in \{1, 2, 3, 4\}$

In the following part, the function of every layer in ANFIS will be explained.

Layer 1: This layer is the input layer, also named as the fuzzification layer. The function of this layer is to obtain fuzzy clustering from input features by using membership function. Each node in this layer represents the values of language variables related to membership functions (such as big and small, high and low). As given in Eq. (1), the output of this layer represents the membership degree of the fuzzy set of input variables transformed into each language variable value.

$$O_i^1 = \mu_{A_i}(x), \quad i=1,2 \quad (1)$$

where x and y are the inputs to node i . Here, fuzzification was done through Gaussian

$$O_i^1 = \mu_{A_i}(x) = e^{-\frac{(x_i - c_i)^2}{2\sigma_i^2}} \quad (2)$$

membership function as it is comprehensible and appropriate to the problem:

where $\{c_i, \sigma_i\}$ is a parameter set. C_i denote the centers of the curve and σ_i determine the

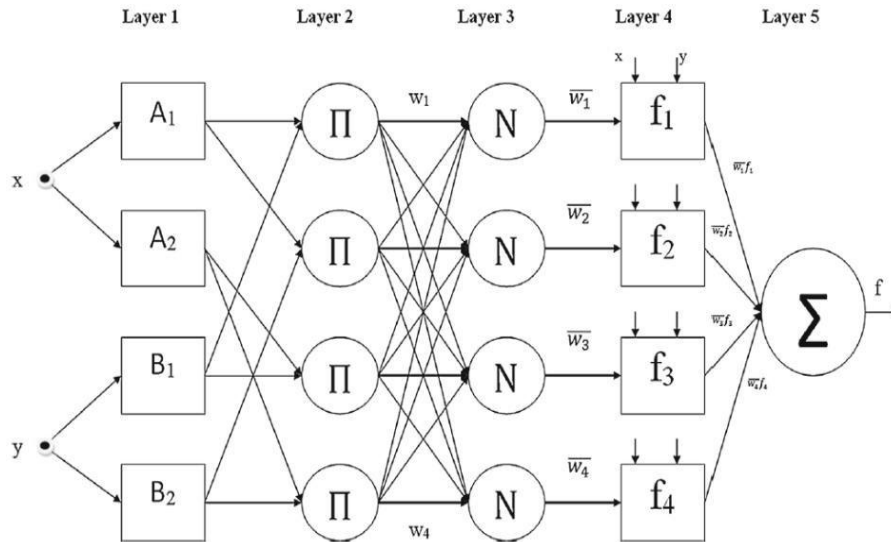


Figure 3.4.
ANFIS
Architecture

membership functions' width. These parameters are called premise parameter.

Layer 2: The function of this layer is to match the premises of fuzzy rules and calculate the firing strength (w_i) of each rule. The output of each model is the product of the input signals.

$$O_i^2 = w_i = \mu_{A_i}(x) \cdot \mu_{B_i}(y), \quad i=1,2 \quad (3)$$

Layer 3: The function of this layer is to implement the normalized calculation. This layer calculates the trigger proportion of rule i in the whole rule base, representing the degree to which this rule is used a`throughout the reasoning process:

$$O_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2 + w_3 + w_4}, \quad i \in \{1,2,3,4\} \quad (4)$$

Layer 4: This layer is the defuzzification layer. In this layer, weighted values of rules are calculated in each node:

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad (5)$$

where w_i is the output of layer 3 and $\{p_i, q_i, r_i\}$ is the consequence parameters set of fuzzy rules.

Layer 5: This layer is the output layer. The overall output of ANFIS is computed by

$$O_i^5 = \sum_i \bar{w}_i f_i = \frac{\sum_{i=1} w_i f_i}{\sum_{i=1} w_i} = \text{overall output} \quad (6)$$

summing all incoming signals:

(Ephzibah *et al.*, 2012).

3.5 TRAINING ANFIS BASED ON GENETIC ALGORITHM

After the fuzzy logic model of ANFIS is established, the next main task is to optimize the parameters via training algorithms. The adjustable parameters in ANFIS are mainly concentrated in layer 1 and layer 4. The parameters of layer 1 are membership function parameters, also known as premise parameters, and the parameters of layer 4 are consequence parameters of fuzzy rules. Inspired by Jang's paper in 1993 [6], conventional ANFIS models often adopt LSE-GD hybrid learning algorithm. To be specific, premise parameters are trained with gradient descent (GD), and consequence parameters are trained with least squares estimation (LSE) method. Nevertheless, after a large number of experiments, it is found that these derivative-based optimization algorithms have a risk of local minimum. Recently, there is a trend that heuristic algorithms are widely used in training ANFIS to achieve better performance, such as genetic algorithm (GA), simulated

annealing algorithm (SAA), particle swarm optimization (PSO), and Artificial immune system (AIS). In this study, GA is applied to optimize the parameters of ANFIS.

GA was proposed by Holland in 1975, which is often applied to search the optimal solution by simulating the process of natural evolution.

Firstly, the algorithm expresses the problem to be solved as chromosome or individual in genetic space by coding process, which was the algorithm's input. And then the algorithm performs genetic processes, such as selection, crossover, and mutation, to search for the global optimum. The flowchart of a basic GA was shown in Fig 2.2.2. The following is the explanation of several crucial steps of GA based on the proposed ANFIS model.

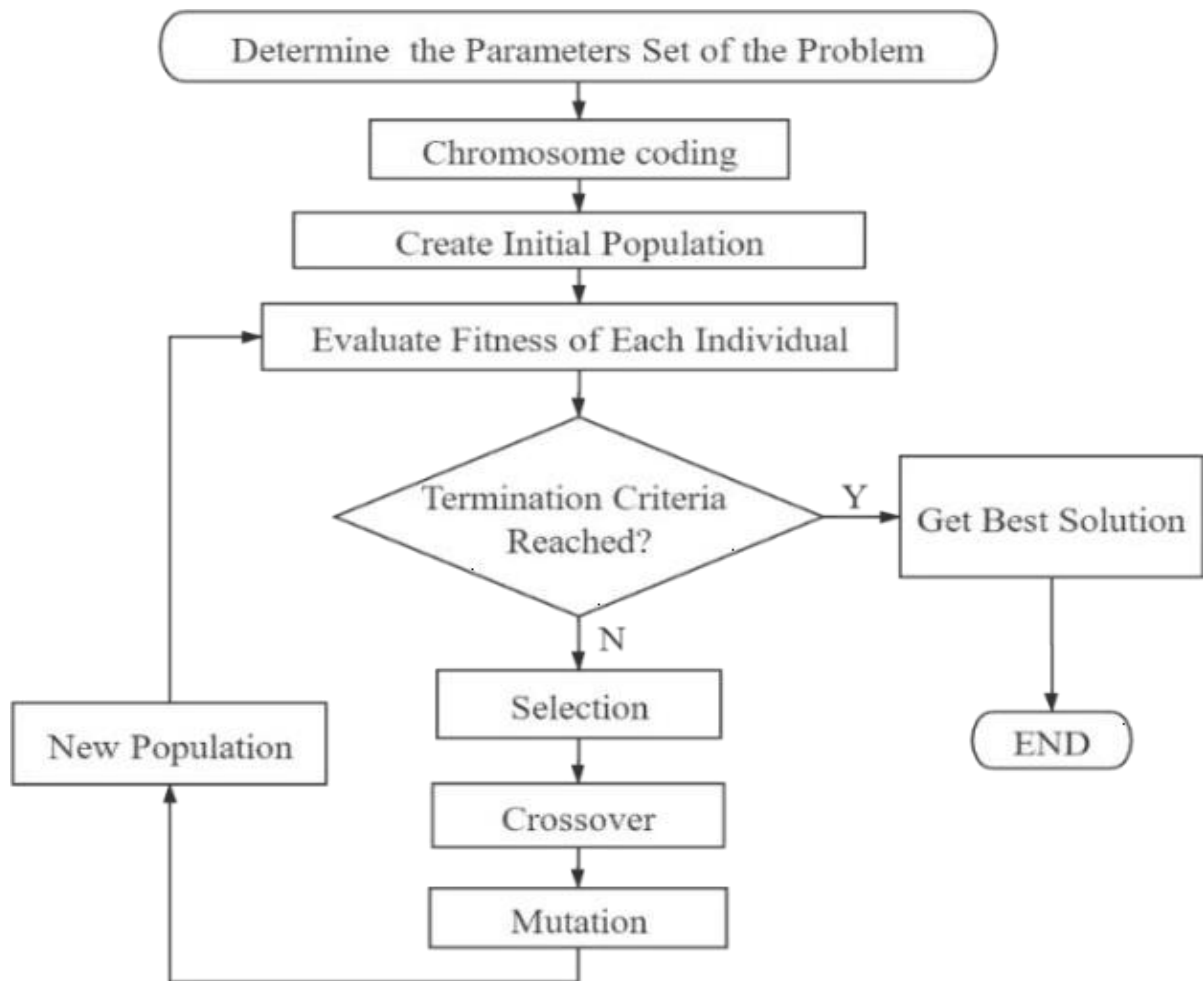
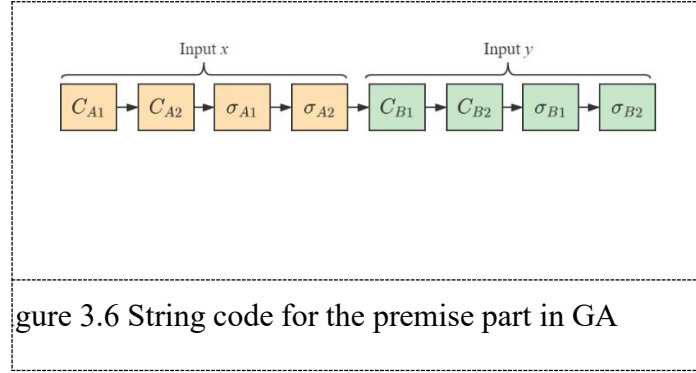


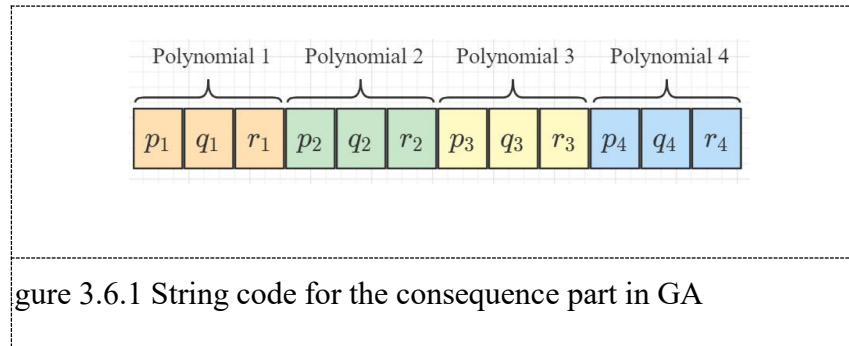
Figure 3.5 Flowchart of a GA

3.6 CHROMOSOME CODING.

As mentioned before, ANFIS has two types of parameters that need to be trained, which are premise and consequence parts. In the premise part, the adjustable parameters are the center (c_i) and the standard deviation (σ_i) of the Gaussian function that is set as the membership function. Therefore, for the premise part of ANFIS structure with two inputs shown in Figure.3.6, the chromosome of the GA can be coded as follows:



For the consequence part, according to the equation (5), the adjustable parameters are the coefficients of each polynomial that are used in defuzzification layer. The chromosome code for the consequent part in GA is:



3.7 FITNESS FUNCTION.

The purpose of the fitness function is to evaluate the quality of the candidate chromosomes (solutions) in the population. In this proposed model, Root Mean Square Error (RMSE) method as Eq. (7) is adopted to calculate the fitness of an individual chromosome. RMSE is a frequently-used method for forecasting technique evaluation. Machine learning models usually minimize RMSE to obtain appropriate parameter values during the training processes. Therefore, the fitness of the chromosome should be inversely proportional to RMSE.

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (d_i - o_i)^2} \quad (7)$$

3.8 SELECTION

The purpose of selection is to select superior individuals from the population with certain probability. Lots of different selection operations have been introduced in previous literatures, such as stochastic uniform, remainder, roulette and tournament. The model of this study uses roulette wheel selection function, also called proportional selection method, which takes into account the fitness of each individual and a certain probability. In brief, candidate chromosomes with better fitness are more likely to be selected as the parents for the next generation. (Feng *et al.*, 2022).

3.8.1 CROSSOVER.

Crossover is the process of genetic recombination between two parents by interchanging chromosome segments in relative positions. There are also lots of ways to implement crossover operation. Based on the chromosome code of the model mentioned above, this study adopts the multi-point crossover method. In brief, the algorithm selects two chromosomes as parents, and randomly set multiple crossover points, then carry out gene exchange, so as to get two different offspring chromosomes. (Feng *et al.*, 2022).

3.8.2 MUTATION.

Once the crossover process is completed, there should be a certain probability of gene mutation to improve the diversity of the population. The mutation probability (p_m) is set very small, generally speaking, $p_m \leq 0.05$. The chromosome of the proposed model is coded by float number, which is suitable for Uniform Mutation. Uniform Mutation refers to replacing the original value of the chosen gene with random numbers conforming to the uniform distribution within a certain range for that gene. This operation is a two-step process. First, each gene in the individual chromosome code string is designated as the mutation point in turn. In the next step, for each mutation point, a

random number is selected uniformly from the value range for that gene by probability pm to replace the original gene value. (Feng *et al.*, 2022).

3.9 DATASET AND FUZZY RULE GENERATION

Dataset Description.

This study used the Cleveland dataset from UCI repository as the experimental dataset, which contains 303 instances with 13 input attributes and one output field which refers to the class label.

In the experiment of this study, the dataset is randomly divided into two sub-sets: training (253 instances) and testing (50 instances). Detailed information of the dataset attributes is shown in

Table 3.7

ATTRIBUTES' INFORMATION OF THE CLEVELAND DATASET

ATTRIBUTE	NUMBER OF MF	DESCRIPTION
Age	3	Age (Young, Mid-Aged, Old)
Sex	2	Sex (0; Female; 1; Male).
Cp	4	Chest pain type (4 types).
Trestbps	4	Resting blood pressure (Low, Medium, High, very high).
Chol	4	Serum cholesterol (Low, Medium, High, very high).
Fbs	2	Fasting blood sugar is greater than 120 mg/dl or not (0; false; 1;true).
Restecg	3	Resting electrocardiographic results (3 types).
Thalach	3	Maximum heart rate (Low, Medium, High).
Exang	2	Exercise induced angina (0; no; 1; yes).

Oldpeak	3	Exercise induced by exercise relative to rest (Low Risk, Terrible).
Slope	3	The slope of the peak exercise ST segment (3 types)
Ca	4	Number of major vessels [0-3] colored by fluoroscopy
Thal	3	The heart status (3 types)

Table 3.7

It is necessary to explain the column “Number of MF” in the table 1, which refers to the number of membership functions needed for a certain attribute in the fuzzification layer of the ANFIS model. For instance, in the original dataset, trestbps is the integer data in, which is fuzzified into four categories as low, medium, high and very high by the calculation of membership function. Therefore, the input “trestbps” needs four nodes in the fuzzification layer.

3.10 FUZZY RULE GENERATION

A reasonable fuzzy rule base directly determines the classification performance of a fuzzy inference system. When ANFIS is implemented with PYTHON, the rule base can be generated automatically or added manually. However, according to Table 2.11, this model has 13 inputs, and each input corresponds to several membership functions, thus the automatically generated rule base will be very large.

Therefore, in order to improve the training speed and the accuracy of the model, it is necessary to extract effective rules. Referring to literature and, and making a proper analysis of the dataset, 87 rules are adopted in this system. An example of these rules is shown below:

Rule 1: If (Age is Young) and (Trestbps is Low) and (Chol is Low) and (Restecg is Normal) and (Thalach is Low) and (Thal is Normal) and (Sex is Male) and (CP is Asymptomatic) and (Fbs is False) and (Exang is No) and (Oldpeak is Low) and (Slope is upsloping) and (Ca is 0) then (RESULT is healthy).

3.11 EXPERIMENTAL RESULTS

To implement the proposed heart disease diagnosis model, PYTHON programming language is utilized. The work is carried out in windows 10 operation system that has Intel Core i5 processor with speed 2.7 GHz and 8GB RAM.

3.12 EVALUATION METRICS

In the experiment, three measure metrics are defined to evaluate the proposed model: specificity, sensitivity and accuracy:

$$\text{Sensitivity} = \text{TP} / (\text{TP} + \text{FN})$$

$$\text{Specificity} = \text{TN} / (\text{TN} + \text{FP})$$

$$\text{Accuracy} = (\text{TN} + \text{TP}) / (\text{TN} + \text{TP} + \text{FN} + \text{FP})$$

TP means True Positive, FP means False Positive, TN means True Negative and FN means False Negative.

3.13 Results

The experimental result for the training set (253 instances) the testing set (50 instances) achieved by the proposed ANFIS-GA model was shown as Table 3.11 A remarkable result can be seen that the

experiment achieved 91.25% accuracy on the testing set.

Table 3.13 Performance evaluation of the proposed model

	Sensitivity	Specificity	Accuracy
Training set	94.35	92.78	93.27
Testing set	91.54	90.32	91.25

3.14 Performance comparison

The following is the performance comparison of several previous heart disease diagnosis system, which all used the Cleveland heart disease dataset. As shown in Table 3.12, the experimental results of

the system proposed in this study was found to be satisfying based on comparison.

Table 3.14 Performance comparison of different methods

Model	Accuracy
Hybrid Neural Network	86.8%
AGAFL	90%
ANN-Fuzzy_AHP [21]	91.1%
Bagging-Fuzzy-GBDT [22]	87%
Proposed Model	91.25%

CHAPTER FOUR

SYSTEM IMPLEMENTATION AND DESIGN

4.1 INTRODUCTION

System implementation involves the actual creation and testing of the model. This chapter aims to construct the proposed model, record the development process, and showcase the results.

4.2 SYSTEM REQUIREMENTS

System requirements are the configurations that the system must know that failure to meet these requirements can affect performance.

4.2.1 HARDWARE REQUIREMENT:

The following hardware are needed to run the model conveniently:

- a. Processor
- b. Ram: 8GB
- c. Memory:120GB
- d. Processor speed: 2.2GHz and above
- e. GPU: A GPU is needed in order to improve the speed of the model as it utilizes multiple cores for processing which the model can use during the training process.

4.2.2 SOFTWARE REQUIREMENT

A software is the part of the computer that cannot be touched. The minimum software requirement needed to run the model are:

- a. Operating System: Windows 10
- b. Development environment: spyder\

4.3 MODEL DEVELOPMENT TOOL

These are the programs that are used by the software developers in creating, debugging, and maintenance of other programs and applications. The major software tools used in the development of the model is given in the next section.

4.3.1 PYTHON

Python is a high-level, general purpose programming language with powerful data types. It is object oriented with a dynamic semantics nature that allows its design ideas to posse code readability and reusability. It allows the use of library which serves as a collection of functions which helps in performing many tasks without necessarily having to rewrite from scratch. Some of the python library utilized in this work include:

NumPy: which stands for Numerical Python, is a fundamental package for scientific computing in Python. It provides support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays. NumPy is essential for data science and forms the backbone of many other Python scientific libraries like SciPy, Matplotlib, and Pandas, making operations like numerical integration, transformation, and visualization more manageable. Its ability to perform complex mathematical operations quickly

Pandas: is a powerful and widely used Python library designed primarily for data manipulation and analysis. It provides highly optimized data structures like Series and Data Frame, making it easy to handle large datasets and perform complex data operations. Pandas simplifies tasks like data cleaning, transformation, and aggregation. It also supports merging and joining of datasets and time series analysis, making it a go-to library for data scientists and analysts working in Python.

Matplotlib: is a popular plotting library for the Python programming language. It provides an object-oriented API for embedding plots into applications using general-purpose GUI toolkits.

Matplotlib is highly versatile and allows for the creation of a wide variety of static, animated, and interactive visualizations in Python. It supports various plot types such as line plots, scatter plots, bar charts, histograms, and more, making it a powerful tool for data visualization.

In Python, the Pickle module is used for serializing and deserializing objects. Serialization (pickling) converts a Python object into a byte stream, and deserialization (unpickling) converts the byte stream back into a Python object. This is useful for saving complex data types like lists, dictionaries, or custom objects to a file or sending them over a network.

4.3.2 Streamlit: is an open-source Python library that simplifies the process of creating and sharing custom web applications for machine learning and data science. It allows you to turn data scripts into interactive web apps in just a few lines of code. Some key features include:

1. **Ease of Use:** Write apps using simple and intuitive Python code.
2. **Interactive Widgets:** Add interactive widgets like sliders, buttons, and text inputs effortlessly.
3. **Real-time Updates:** Automatically updates apps in real-time when code or data changes.
4. **Integration:** Easily integrates with popular data science libraries like Pandas, NumPy, Matplotlib, and more.

Streamlit is ideal for quickly prototyping data analysis tools and sharing them with others without needing extensive web development knowledge.

4.3.1.4 ANACONDA NAVIGATOR

is a desktop graphical user interface (GUI) included in the Anaconda distribution. It allows users to manage their Python and R environments and packages with ease. Key features include:

1. **Environment Management:** Create, manage, and switch between different virtual environments.

2. Package Management: Install, update, and remove packages from various repositories.
3. Integrated Tools: Launch popular applications like Jupyter Notebook, Spyder, and RStudio directly from the interface.
4. User-Friendly: Provides a simple and intuitive interface for users who prefer not to use command-line tools.

Anaconda Navigator simplifies the process of managing data science and machine learning projects.

4.4 SYSTEM TESTING

System testing is used to test for the performance of the model. Testing the model is generally about making use of the model with the test dataset and giving a comparison with the model's performance in terms of accuracy, recall, precision.

TABLE 4.1

DATASET

age	sex	chest	trestbps	cholesterol	fastingbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
57	0	0	120	354	0	1	163	1	0.6	2	0	2	1
57	1	0	140	192	0	1	148	0	0.4	1	0	1	1
56	0	1	140	294	0	0	153	0	1.3	1	0	2	1
44	1	1	120	263	0	1	173	0	0	2	0	3	1
52	1	2	172	199	1	1	162	0	0.5	2	0	3	1
57	1	2	150	168	0	1	174	0	1.6	2	0	2	1
54	1	0	140	239	0	1	160	0	1.2	2	0	2	1
48	0	2	130	275	0	1	139	0	0.2	2	0	2	1
49	1	1	130	266	0	1	171	0	0.6	2	0	2	1
64	1	3	110	211	0	0	144	1	1.8	1	0	2	1
58	0	3	150	283	1	0	162	0	1	2	0	2	1
50	0	2	120	219	0	1	158	0	1.6	1	0	2	1
58	0	2	120	340	0	1	172	0	0	2	0	2	1
66	0	3	150	226	0	1	114	0	2.6	0	0	2	1
43	1	0	150	247	0	1	171	0	1.5	2	0	2	1
69	0	3	140	239	0	1	151	0	1.8	2	2	2	1
59	1	0	135	234	0	1	161	0	0.5	1	0	3	1
44	1	2	130	233	0	1	179	1	0.4	2	0	2	1
42	1	0	140	226	0	1	178	0	0	2	0	2	1
61	1	2	150	243	1	1	137	1	1	1	0	2	1
40	1	3	140	199	0	1	178	1	1.4	2	0	3	1
71	0	1	160	302	0	1	162	0	0.4	2	2	2	1

59	1	2	150	212	1	1	157	0	1.6	2	0	2	1
51	1	2	110	175	0	1	123	0	0.6	2	0	2	1
65	0	2	140	417	1	0	157	0	0.8	2	1	2	1
53	1	2	130	197	1	0	152	0	1.2	0	0	2	1
41	0	1	105	198	0	1	168	0	0	2	1	2	1
65	1	0	120	177	0	1	140	0	0.4	2	0	3	1
44	1	1	130	219	0	0	188	0	0	2	0	2	1
54	1	2	125	273	0	0	152	0	0.5	0	1	2	1
51	1	3	125	213	0	0	125	1	1.4	2	1	2	1
46	0	2	142	177	0	0	160	1	1.4	0	0	2	1
54	0	2	135	304	1	1	170	0	0	2	0	2	1
54	1	2	150	232	0	0	165	0	1.6	2	0	3	1
65	0	2	155	269	0	1	148	0	0.8	2	0	2	1
65	0	2	160	360	0	0	151	0	0.8	2	0	2	1
51	0	2	140	308	0	0	142	0	1.5	2	1	2	1
48	1	1	130	245	0	0	180	0	0.2	1	0	2	1
45	1	0	104	208	0	0	148	1	3	1	0	2	1
53	0	0	130	264	0	0	143	0	0.4	1	0	2	1
39	1	2	140	321	0	0	182	0	0	2	0	2	1
52	1	1	120	325	0	1	172	0	0.2	2	0	2	1
44	1	2	140	235	0	0	180	0	0	2	0	2	1
47	1	2	138	257	0	0	156	0	0	2	0	2	1
53	0	2	128	216	0	0	115	0	0	2	0	0	1

```
else:
```

```
    parkinsons_diagnosis = "The person does not have Parkinson's disease"
```

```
st.success(parkinsons_diagnosis)
```

```
C:\Windows\system32\cmd.exe - streamlit run "C:\Users\USER\Desktop\multiple-disease-prediction-streamlit-app\heartdiseasedig.py"

(machineLearning) C:\Users\USER>streamlit run "C:\Users\USER\Desktop\multiple-disease-prediction-streamlit-app\heartdiseasedig.py"

You can now view your Streamlit app in your browser.

Local URL: http://localhost:8501
Network URL: http://10.10.2.112:8501

C:\ProgramData\anaconda3\envs\machineLearning\Lib\site-packages\sklearn\base.py:376: InconsistentVersionWarning: Trying to unpickle estimator SVC from version 1.0.2 when using version 1.4.1.post1. This might lead to breaking code or invalid results. Use at your own risk. For more info please refer to: https://scikit-learn.org/stable/model_persistence.html#security-maintainability-limitations
  warnings.warn(
C:\ProgramData\anaconda3\envs\machineLearning\Lib\site-packages\sklearn\base.py:376: InconsistentVersionWarning: Trying to unpickle estimator LogisticRegression from version 1.0.2 when using version 1.4.1.post1. This might lead to breaking code or invalid results. Use at your own risk. For more info please refer to: https://scikit-learn.org/stable/model_persistence.html#security-maintainability-limitations
  warnings.warn(
C:\ProgramData\anaconda3\envs\machineLearning\Lib\site-packages\sklearn\base.py:376: InconsistentVersionWarning: Trying to unpickle estimator SVC from version 1.0.2 when using version 1.4.1.post1. This might lead to breaking code or invalid results. Use at your own risk. For more info please refer to: https://scikit-learn.org/stable/model_persistence.html#security-maintainability-limitations
  warnings.warn(
C:\ProgramData\anaconda3\envs\machineLearning\Lib\site-packages\sklearn\base.py:376: InconsistentVersionWarning: Trying to unpickle estimator LogisticRegression from version 1.0.2 when using version 1.4.1.post1. This might lead to breaking code or invalid results. Use at your own risk. For more info please refer to: https://scikit-learn.org/stable/model_persistence.html#security-maintainability-limitations
  warnings.warn(
```

4.6 TERMINAL EXECUTION

🔗 Heart Disease Prediction using ML

Age	Sex	Chest Pain types
<input type="text"/>	<input type="text"/>	<input type="text"/>
Resting Blood Pressure	Serum Cholesterol in mg/dl	Fasting Blood Sugar > 120 mg/dl
<input type="text"/>	<input type="text"/>	<input type="text"/>
Resting Electrocardiographic results	Maximum Heart Rate achieved	Exercise Induced Angina
<input type="text"/>	<input type="text"/>	<input type="text"/>
ST depression induced by exercise	Slope of the peak exercise ST segment	Major vessels colored by flourosopy
<input type="text"/>	<input type="text"/>	<input type="text"/>
thal: 0 = normal; 1 = fixed defect; 2 = reversable defect		
<input type="text"/>		
<input type="button" value="Heart Disease Test Result"/>		
<div style="background-color: #e0ffe0; height: 20px; width: 100%;"></div>		

4.7 PYTHON CODE OUTPUT

CHAPTER FIVE

CONCLUSION, RECONMENDATION AND FUTURE WORKS

5.1 CONCLUSION

The research demonstrates the effectiveness of fuzzy logic and genetic algorithms in improving the accuracy and adaptability of heart disease diagnosis over traditional methods. Adaptive Neuro-Fuzzy Inference Systems (ANFIS), particularly when enhanced with genetic algorithms, are highly accurate in predicting heart disease due to their ability to handle complex clinical interactions. Furthermore, the use of explainable AI techniques in heart attack prediction increases the transparency and reliability of predictive models, enabling healthcare professionals to comprehend and validate automated decisions effectively.

5.2 RECOMMENDATIONS

Medical practitioners are encouraged to integrate advanced AI tools such as ANFIS and fuzzy logic systems into clinical practices to improve diagnostic precision and efficiency, with suggested training for effective model output interpretation. Researchers should enhance the explainability of AI models and include more diverse datasets to boost model robustness across different populations. For technology developers, the focus should be on creating user-friendly interfaces and enhancing the scalability and speed of AI systems to support real-time diagnostics.

5.3 FUTURE WORKS SUMMARY

Future research should aim to incorporate more diverse and multi-center clinical data to enhance the generalizability of predictive models. There should also be an exploration into integrating AI systems with emerging technologies like blockchain and IoT to boost their capabilities. Additionally, conducting longitudinal studies is essential to evaluate the long-term accuracy and

reliability of AI-based diagnostic tools in real-world settings, which will provide insights into their sustained impact across different healthcare scenarios.

REFERENCE

- Abidin, B., Dom, R. M., Rahman, A. R. A., & Bakar, R. A. (Year not provided). Use of Fuzzy Neural Network to Predict Coronary Heart Disease in a Malaysian Sample. Proceedings of the 8th WSEAS International Conference on Telecommunications and Informatics, 76-80. ISBN: 978-960-474-084-0.
- Abushariah, M. A. M., Alqudah, A. A. M., Adwan, O. Y., & Yousef, R. M. M. (2014). Automatic heart disease diagnosis system based on artificial neural network (ANN) and adaptive neuro-fuzzy inference systems (ANFIS) approaches. *Journal of Software Engineering and Applications*, 1055-1064. <https://doi.org/10.4236/jsea.2014.712093>
- Aghamohammadi M et al., (2019). Predicting Heart Attack through Explainable Artificial Intelligence. *Predicting Heart Attack through Explainable Artificial Intelligence*. https://doi.org/10.1007/978-3-030-22741-8_45
- Ahmad, R., Idriss, U. S., & Umar, M. (2018). Using artificial neural networks to diagnose heart disease. **International Journal of Computer Applications*, 182*(19). <https://doi.org/10.5120/ijca2018917938>
- Ali, M. L., Sadi, M. S., & Goni, M. O. (2024). Diagnosis of heart diseases: A fuzzy-logic-based approach. **PLOS ONE*, 19*(2), e0293112. <https://doi.org/10.1371/journal.pone.0293112>.
- Bhagat, R. K., Yadav, A., Rajoria, Y. K., Raj, S., & Boadh, R. (2022). Study of fuzzy and artificial neural network (ANN) based techniques to diagnose heart disease. **Journal of Pharmaceutical Negative Results*, 13*(Special Issue 5), 1023-1033. <https://doi.org/10.47750/pnr.2022.13.S05.161>
- Dehdar Karsidani, S., Farhadian, M., Mahjub, H., & Mozayanimonfared, A. (2022). Intelligent prediction of major adverse cardiovascular events (MACCE) following percutaneous

- coronary intervention using ANFIS-PSO model. *BMC Cardiovascular Disorders*, 22, 389.
<https://doi.org/10.1186/s12872-022-02825-0>
- Ephzibah et al., (2012). Framing fuzzy rules using support sets for effective heart disease diagnosis. *International Journal of Fuzzy Logic Systems*,
<https://doi.org/10.5121/ijfls.2012.2102>
- Feng J et al., (2021). An Intelligent System for Heart Disease Prediction using Adaptive Neuro-Fuzzy Inference Systems and Genetic Algorithm. *Journal of Physics: Conference Series*, 2010, 012172. IOP Publishing. <https://doi.org/10.1088/1742-6596/2010/1/012172>
- Keikhosrokiani, P., Naidu, A. B., Fadilah, S. I., Manickam, S., & Li, Z. (2023). Heartbeat sound classification using a hybrid adaptive neuro-fuzzy inferences system (ANFIS) and artificial bee colony. *Digital Healt*. <https://doi.org/10.1177/20552076221150741>
- Marateb, H. R., & Goudarzi, S. (2015). A noninvasive method for coronary artery diseases diagnosis using a clinically-interpretable fuzzy rule-based system. *Journal of Research in Medical Sciences*, 20*(3), 214-223. Retrieved from <https://www.researchgate.net/publication/328357904>.
- Rajabzadeh, Z *et al.*, (2023). Assessment of the Performances of Adaptive Neuro-Fuzzy Inference System and Two Statistical Methods for Diagnosing Coronary Artery Disease. *Medical Journal of the Islamic Republic of Iran*. <https://doi.org/10.47176/mjiri.37.46>
- Rajinikanth, N., & Pavithira, L. (2019). Heart Disease Prediction and Classification Using Optimization Based Adaptive Neuro Fuzzy Inference System. *Journal of Emerging Technologies and Innovative Research*. Retrieved from <http://www.jetir.org>.
- Rezaei-Hachesu, P., Dehghani-Soufi, M., Khara, R., Moftian, N., & Samad-Soltani, T. (2020). A fuzzy mobile decision support system for diagnosing of the angiographic status of heart

disease. Engineering and Applied Science Research, 47(2), 175-181. <https://www.tci-thaijo.org/index.php/easr/index>. DOI: 10.14456/easr.2020.19

Sreedran, S., Ibrahim, N., Sari, S., Hong Seng, G., & Shanta, S. (2023). Fuzzy rules-based prediction of heart conditions system. *Indonesian Journal of Electrical Engineering and Computer Science, 32*(1), 530-536. <https://doi.org/10.11591/ijeecs.v32.i1.pp530-536>

Taylan, O., Alkabaa, A. S., Alqabbaa, H. S., Pamukçu, E., & Leiva, V. (2023). Early prediction in classification of cardiovascular diseases with machine learning, neuro-fuzzy and statistical methods. *Biology, 12*(1), 117. <https://doi.org/10.3390/biology12010117>

Warhade, S. R., & Hore, U. W. (2017). Intelligent prediction of heart disease diagnosis using ANFIS classification model. *International Journal of Science and Research (IJSR). Retrieved from <http://www.ijsr.net>

Ziasabounchi, N., & Askerzade, I. (2014). ANFIS based classification model for heart disease prediction. *International Journal of Engineering & Computer Science IJECS-IJENS, 14. https://www.researchgate.net/publication/261727412_ANFIS_based_classification_model_for_heart_disease_prediction.

APPENDIX

(SOURCE CODE)

4.5 PYTHON CODE

```
import os

import pickle

import streamlit as st

from streamlit_option_menu import option_menu

# Set page configuration
st.set_page_config(page_title="Health Assistant",
                   layout="wide",
                   page_icon="")

# getting the working directory of the main.py
working_dir = os.path.dirname(os.path.abspath(__file__))

# loading the saved models

diabetes_model = pickle.load(open(f'{working_dir}/saved_models/diabetes_model.sav', 'rb'))

heart_disease_model = pickle.load(open(f'{working_dir}/saved_models/heart_disease_model.sav', 'rb'))
```

```

parkinsons_model = pickle.load(open(f'{working_dir}/saved_models/parkinsons_model.sav',
'rb'))

# sidebar for navigation
with st.sidebar:

    selected = option_menu('Multiple Disease Prediction System',

                            ['Diabetes Prediction',
                             'Heart Disease Prediction',
                             'Parkinsons Prediction'],
                            menu_icon='hospital-fill',
                            icons=['activity', 'heart', 'person'],
                            default_index=0)

# Diabetes Prediction Page
if selected == 'Diabetes Prediction':

    # page title
    st.title('Diabetes Prediction using ML')

    # getting the input data from the user
    col1, col2, col3 = st.columns(3)

    with col1:

```

```
Pregnancies = st.text_input('Number of Pregnancies')

with col2:

    Glucose = st.text_input('Glucose Level')

with col3:

    BloodPressure = st.text_input('Blood Pressure value')

with col1:

    SkinThickness = st.text_input('Skin Thickness value')

with col2:

    Insulin = st.text_input('Insulin Level')

with col3:

    BMI = st.text_input('BMI value')

with col1:

    DiabetesPedigreeFunction = st.text_input('Diabetes Pedigree Function value')

with col2:

    Age = st.text_input('Age of the Person')

# code for Prediction
```

```

diab_diagnosis = "

# creating a button for Prediction

if st.button('Diabetes Test Result'):

    user_input = [Pregnancies, Glucose, BloodPressure, SkinThickness, Insulin,
                  BMI, DiabetesPedigreeFunction, Age]

    user_input = [float(x) for x in user_input]

    diab_prediction = diabetes_model.predict([user_input])

    if diab_prediction[0] == 1:
        diab_diagnosis = 'The person is diabetic'
    else:
        diab_diagnosis = 'The person is not diabetic'

    st.success(diab_diagnosis)

# Heart Disease Prediction Page

if selected == 'Heart Disease Prediction':

    # page title

    st.title('Heart Disease Prediction using ML')

```

```
col1, col2, col3 = st.columns(3)

with col1:
    age = st.text_input('Age')

with col2:
    sex = st.text_input('Sex')

with col3:
    cp = st.text_input('Chest Pain types')

with col1:
    trestbps = st.text_input('Resting Blood Pressure')

with col2:
    chol = st.text_input('Serum Cholestorol in mg/dl')

with col3:
    fbs = st.text_input('Fasting Blood Sugar > 120 mg/dl')

with col1:
    restecg = st.text_input('Resting Electrocardiographic results')

with col2:
```

```

thalach = st.text_input('Maximum Heart Rate achieved')

with col3:

    exang = st.text_input('Exercise Induced Angina')

with col1:

    oldpeak = st.text_input('ST depression induced by exercise')

with col2:

    slope = st.text_input('Slope of the peak exercise ST segment')

with col3:

    ca = st.text_input('Major vessels colored by flourosopy')

with col1:

    thal = st.text_input('thal: 0 = normal; 1 = fixed defect; 2 = reversable defect')

# code for Prediction

heart_diagnosis = "

# creating a button for Prediction

if st.button('Heart Disease Test Result'):

```

```

user_input = [age, sex, cp, trestbps, chol, fbs, restecg, thalach, exang, oldpeak, slope, ca,
thal]

user_input = [float(x) for x in user_input]

heart_prediction = heart_disease_model.predict([user_input])

if heart_prediction[0] == 1:
    heart_diagnosis = 'The person is having heart disease'
else:
    heart_diagnosis = 'The person does not have any heart disease'

st.success(heart_diagnosis)

# Parkinson's Prediction Page
if selected == "Parkinsons Prediction":

    # page title
    st.title("Parkinson's Disease Prediction using ML")

    col1, col2, col3, col4, col5 = st.columns(5)

    with col1:
        fo = st.text_input('MDVP:Fo(Hz)')

```

```
with col2:
```

```
    fhi = st.text_input('MDVP:Fhi(Hz)')
```

```
with col3:
```

```
    flo = st.text_input('MDVP:Flo(Hz)')
```

```
with col4:
```

```
    Jitter_percent = st.text_input('MDVP:Jitter(%)')
```

```
with col5:
```

```
    Jitter_Abs = st.text_input('MDVP:Jitter(Abs)')
```

```
with col1:
```

```
    RAP = st.text_input('MDVP:RAP')
```

```
with col2:
```

```
    PPQ = st.text_input('MDVP:PPQ')
```

```
with col3:
```

```
    DDP = st.text_input('Jitter:DDP')
```

```
with col4:
```

```
    Shimmer = st.text_input('MDVP:Shimmer')
```

```
with col5:
```

```
Shimmer_dB = st.text_input('MDVP:Shimmer(dB)')

with col1:
    APQ3 = st.text_input('Shimmer:APQ3')

with col2:
    APQ5 = st.text_input('Shimmer:APQ5')

with col3:
    APQ = st.text_input('MDVP:APQ')

with col4:
    DDA = st.text_input('Shimmer:DDA')

with col5:
    NHR = st.text_input('NHR')

with col1:
    HNR = st.text_input('HNR')

with col2:
    RPDE = st.text_input('RPDE')

with col3:
    DFA = st.text_input('DFA')
```

```

with col4:

    spread1 = st.text_input('spread1')

with col5:

    spread2 = st.text_input('spread2')

with col1:

    D2 = st.text_input('D2')

with col2:

    PPE = st.text_input('PPE')

# code for Prediction

parkinsons_diagnosis = "

# creating a button for Prediction

if st.button("Parkinson's Test Result"):

    user_input = [fo, fhi, flo, Jitter_percent, Jitter_Abs,

                  RAP, PPQ, DDP, Shimmer, Shimmer_dB, APQ3, APQ5,

                  APQ, DDA, NHR, HNR, RPDE, DFA, spread1, spread2, D2, PPE]

    user_input = [float(x) for x in user_input]

```

```
parkinsons_prediction = parkinsons_model.predict([user_input])
```

```
if parkinsons_prediction[0] == 1:
```

```
    parkinsons_diagnosis = "The person has Parkinson's disease"
```