CHAPTER 1
INTRODUCTION

The advancements of computer technology are playing an important role in diagnosis of medical systems. Medical diagnosis is full of uncertainty and dynamically changed according to the situations. Now a day the use of computer technology is essential in every filed and medical diagnosis area is not an exception. The diagnostic decision depends upon experience, expertise and use of proper method with powerful logical reasoning ability. We know very well that these fields, in which the computers are used, have very high complexity and irregularity and the use of expert systems such as fuzzy logic, artificial network and genetic algorithm have been developed. Expert Systems is an intelligent computer-based decision-making tool that uses data sets and rules to solve real problems based on the knowledge gained by one or more experts in specific areas. The medical diagnostic system is a system that can diagnose any disease by controlling the symptoms. The diagnosis of human diseases is one of the most complex and difficult processes and requires a high level of competence. The Fuzzy Expert System is one of the best systems to diagnose medical conditions, because every disease diagnosis involves so many uncertainties and a fuzzy logic is the best way to handle uncertainties. Despite the limitations due to information, education and other reasons, these systems are widely recognized in medical institutions.

1.1 Introduction to Fuzzy

Fuzzy logic (FL) is a multi-valued logic that has been used to solve many complex challenges including clinical diagnostics. FL handles approximate values in place of fixed and precise values. Professor Lotfi A. Zadeh first introduced the terms “fuzzy sets” and “fuzzy logic” in the mid-1960’s [1-2].

According to Zadeh, fuzzy logic is an addition of the classic logic. Classic logic is based on boolean logic, where the information is either true or false. In classic logic the membership of a component belonging to a set is represented by 0 if it does not belong
to the set and 1 if it is in the set, i.e. \{0, 1\}. On the other hand, in fuzzy logic this set is extended to the interval of [0, 1].

1.1.1 Membership Function

A membership function (MF) is a distribution that maps each and every point in the input space (i.e., universe of discourse which represents the set of the entities) to a membership value between 0 and 1. There are different types of membership functions of fuzzy set such as triangular membership function, trapezoidal membership function, Gaussian membership function, and etc.[3]. The types of MF depend on the concept that is being represented, and the context of it is use. This study used triangular and trapezoidal membership functions.

In triangular membership function, the curve is a vector function \(x\) to be determined by three scalars \(a, b,\) and \(c\). In fig. 1.1, a triangular membership function is illustrated [3].

![Figure 1.1: Triangular Membership Function](image1)

In trapezoidal membership function, the curve is a vector function \(x\), determined by four scalars \(a, b, c\) and \(d\). In the following figure (1.2), a trapezoidal membership function is illustrated [3-4].

![Figure 1.2: Trapezoidal Membership Function](image2)
Membership function can be the combination of both of them. For example, in the following figure (1.3), the triangular and the trapezoidal membership functions (MF) are illustrated:

![Membership functions](image)

**Figure 1.3: A Basic fuzzy Set of Triangular and Trapezoidal Membership Function**

However, Gaussian, Sigmoid and other types of linear functions can also be applied to characterize the fuzzy sets. Non-linear functions can also be used but they will cause additional computational complexity to the algorithm [5].

### 1.1.2 Types of FuzzySets

There are mainly two types of fuzzy sets: type-1 fuzzy sets (T1FS) and type-2 fuzzy sets (T2FS). T1FS were first introduced by L.A. Zadeh in 1965 [1]. However, type-1 fuzzy sets failed to model uncertainty properly. Uncertainty indicates the degree of truth of the value in attribute. For example, the age of John is 36 right now, might be 80% true. The issues with T1FS led to the introduction of T2FS. In order to model uncertainty and imprecision in a superior way, type-2 fuzzy set was initially presented by Lotfi Zadeh and the concepts were presented by Mendel and Liang [6]. In case of T2FS, the degree of membership is type-1 fuzzy set. The following example will explain more about the idea of type-1 and type-2 fuzzy set:
In figure 1.4, temperature $x = 27^\circ C$ is hot to the degree of membership 0.4 (i.e. $\mu_{Temperature}(27) = 0.4$). In this example, there is vagueness but no uncertainty. In vagueness, the elements are vague and represented by linguistic terms (which is hot in this example). To model uncertainty, it is required to know the temperature is hot to a degree about 0.4. In order to model uncertainty properly, type-2 fuzzy set (T2FS) is introduced. The example of type-1 fuzzy set shown in figure 1.4 can be replaced by the following figure (1.5) using type-2 fuzzy set.

Figure 1.5 shows that for type-2 fuzzy set, the membership function is an interval i.e. type-1 fuzzy set. For the above example of temperature $x = 27^\circ C$, the membership function for $x = 27^\circ C$ is $[0.4, 1]$.

In type-2 fuzzy set, the membership function is three dimensional. The third dimension is called the footprint of uncertainty (FOU) and it is the membership value at each point of the two dimensional domain. FOU is represented by lower bound and upper bound..
membership function, both of which are type-1 fuzzy set. For the above example, \( x = 27^\circ C \) has lower MF 0.4 and upper MF 1.

### 1.1.3 Operations on Fuzzy Sets

There are mainly three operations on fuzzy sets, which are complement, intersection and union. Let \( A \) and \( B \) be two fuzzy sets defined on the universe of discourse \( X \) to the interval \([0,1]\). A fuzzy set \( A \) is defined by its membership function \( \mu(A) \) and for the fuzzy set \( B \), which is defined by its membership function \( \mu(B) \) over \( X \). The function-theoretical operations of union, intersection and complements are defined as follows:

**Union:**

\[
\mu_{A \cup B}(x) = \mu_A(x) \lor \mu_B(x) = \max(\mu_A(x), \mu_B(x))
\]

**Intersection:**

\[
\mu_{A \cap B}(x) = \mu_A(x) \land \mu_B(x) = \min(\mu_A(x), \mu_B(x))
\]

**Complement:**

\[
\mu_{\neg A}(x) = 1 - \mu_A(x)
\]

The associativity and commutativity of minimum and maximum functions are defined by the following [7]:

\[
\max(x, \max(y, z)) = \max(\max(x, y), z)
\]

\[
\min(x, \min(y, z)) = \min(\min(x, y), z)
\]

![Fuzzy Sets Operations](image)

**Figure 1.6: Fuzzy Sets Operations [4]**

Figure 1.6 shows the operations of fuzzy sets i.e., union, intersection and complement. Any fuzzy set \( A \) defined on the universe \( X \) is a subset of the universe. Also, by definition...
the null set has membership 0 and \( x \) in \( X \) has membership 1. Note that the null set and the whole set are not fuzzy sets.

1.2 Linguistic Variables and Linguistic Values

A linguistic variable is a fuzzy variable [8]. In mathematics, variables deal with numerical values, whereas in fuzzy logic applications, non-numerical linguistic variables are usually used to assist the progress of the expression of rules and facts. In linguistic variables, the values come from natural language or artificial language such as words or sentences [9]. For example, “Age is Young” indicates the linguistic variable “Age” accepts “Young” which is linguistic value.

1.2.1 Fuzzy IF-THEN Rules

The rules are the heart of the fuzzy inference system. After defining the linguistic variables and values, the rules of the fuzzy system can be formulated. The rules are used to map fuzzy inputs to fuzzy outputs. Fuzzy rules have three parts: antecedent, proposition and consequence(s). One antecedent may have more than one of the (AND) or (OR) operators. The fuzzy IF-THEN rule looks like the following:

Rule: 1 IF \( x \) is \( A_1 \) OR \( y \) is \( B_1 \) THEN \( z \) is \( C_1 \).
Rule: 2 IF \( x \) is \( A_2 \) AND \( y \) is \( B_2 \) THEN \( z \) is \( C_2 \).
Rule: 3 IF \( x \) is \( A_3 \) THEN \( z \) is \( C_3 \).

Where \( A, B \) and \( C \) are the linguistic values and \( x, y, \) and \( z \) are the linguistic variables.

1.2.2 Fuzzy Inference System

A fuzzy inference system (FIS) uses fuzzy set theory in order to map input to output. All information is involved in the FIS process, i.e. membership functions, logical operators and IF-THEN rules. A sample FIS that includes four functions illustrated in figure 1.7.
There are two types of FIS [93] i.e. the Mamdani [8] and the Sugeno [11]. Such FIS are used in many researches such as expert system and decision support system.

**Mamdani Type FIS**

If-then rules are applied in mamdani based FIS for inputs and output. For example: IF X is Negative Big AND Y is NegativeSmall THEN Z is Zero.

**Sugeno Type FIS**

Sugeno type systems are used to model any inference system in which membership functions are linear or constant. This fuzzy inference system was introduced in 1985. It is also called Takagi-Sugeno-Kang. The functions of belonging to the Sugeno (z) output are linear or constant. A typical rule in a fuzzy model of Sugeno is:

If Input 1 = x and Input 2 = y, then Output is z = ax + by + c

For a zero-order Sugeno model, the output level z is a constant (a=b =0).

Sugeno and Mamdani FIS can be used for similar tasks. The rule base and the fuzzification remain the same for the variables. There are several defuzzifiers that can be selected for a FIS Mamdani. These editors also have similar results in a FIS Sugeno. There is a certain overlap between the two types of system. Mamdani FIS is used more often. It is used for decision support applications because of its intuitive and interpretable nature. The consequences of the rules in a FIS Sugeno do not have a direct
DESIGN AND DEVELOPMENT OF FUZZY EXPERT SYSTEM FOR DISEASE DIAGNOSIS

semantic average. This means that they are not linguistic concepts. Even this interpretability is partially lost. The consequences of Sugeno's FIS rules can have many parameters per rule. So, Sugeno FIS translates into more degrees of freedom than Mamdani FIS.

Thus, it offers greater flexibility. Many parameters can be used in the consequences of the rules of a FIS Sugeno. A zero order FIS Sugeno can reasonably approach a Mamdani FIS. From a mathematical point of view, a FIS Sugeno is more effective than a Mamdani FIS. This because; Sugén FIS does not involve an intense process of defuzzificazione IT. In addition, a FIS Sugeno always produces consistent surfaces. The continuity of the output interface is very important. Any existence of discontinuity translates into similar inputs that produce substantially different results. This will be an undesirable situation from the point of view of control / surveillance. Because of the continuous structure of the output functions, a FIS Sugeno is also better and more suitable for the functional analysis of a Mamdani FIS.

This thesis uses the Muddani inference method. This is the most used fuzzy method for its simple structure of "min-max" or "AND-OR" operations. The Mamdani method was proposed in 1975 by Professor Ebrahim Mamdani at the University of London [8]. Muddani's fuzzy deduction process includes four phases: fuzzification; Evaluation rule; Power unit (s); and the defuzzification described below[12-13]:

**Step 1: Fuzzification**

In fuzzification process the crisp input values are transformed into the grades of membership function (MF) for linguistic values of fuzzy sets. The membership function provides a grade for each linguistic value.

**Step 2: Rule Evaluation**

After successfully defining the input and output variables, and the corresponding MFs, it is necessary to design the rule-base of the fuzzy knowledge-base. The rule-base of FIS design is composed of IF < antecedents > THEN < conclusion > rules. These rules are then transformed from an input to an output, based on MFs that inform the projected outcomes. The total number of rules depends on the total number of linguistic variables and MFs. In Mamdani, the AND operator is applied on each rule for rule evaluation.
Step 3: Aggregate output(s)
After evaluating all the rules, the rules must be grouped in a particular approach to make a decision. The aggregation method is used to group the fuzzy output set after evaluating the rules. In Mamdani, the OR operator is used to aggregate fuzzy output sets. After aggregation, the final result is a single fuzzy set.

Step 4: Defuzzification
The defuzzification is placed after all other processes of fuzzy inference. This method is used to generate a precise number from the fuzzy set of a single output obtained by aggregating the rules in step 3. Several methods are used for defuzzification, such as the value of centroid (and the center of gravity or COG) [12], the bisector (bisector) [12], the average value of the maximum value (MOM) [12], the smallest (absolute) of the maximum value (SOM) [12] and finally the largest (absolute) value of the maximum (LOM) [12]. The centroid method is the most common defuzzification method [12]. The centroid defuzzification method is used to determine the point indicating the center of gravity of the fuzzy set.

In the case of a FIS project, all the previous steps must be taken into consideration. Elements of the conceptual schema include tables, views, constraints, domain definitions, and other constructs that describe the schema. For the definition of a fuzzy database system it can be described how the database system now uses fuzzy logic to support vague, inaccurate and uncertain information. In other words, the fuzzy database is called a fuzzy database [3], [14].

1.3 Fuzzy Inference System in General Application

This section describes a simple fuzzy inference system for the Hotel Advisory System (HAS) [93]. Nagi et al. [15] have developed an expert blurred hotel selection system called Hotel Advisory System (HAS). The system is divided into three modules. The modules are the hotel fuzzy search form, hotel information and the hotel's virtual tour module. Each form is used for a specific activity. A user interface is designed so that the user can enter data and return information to the user. The system used a fuzzy inference system to provide linguistic terms such as cheap, moderate, expensive for the price of
the hotel and the fuzzy rules of the system were used to determine the cost of staying in the hotel. This fuzzy expert system is more convenient and easier for the user to choose the hotel based on their demand such as cheaper hotel, expensive hotel. The system was tested by the potential users and hotel experts. It can be used to improve the operations by reducing the cost of enquires, and providing quick information about the hotel search.

1.4 Related Works of Fuzzy Inference System in Medical Diagnostic Systems

In this section different contributions of researchers using fuzzy inference system (FIS) is discussed [93]. Many research works have been developed using fuzzy logic in the diagnosis of various diseases using MATLAB simulation platform. Adeli and Neshat [16] proposed an expert system for diagnosing heart disease using fuzzy logic. In this research, in order to get fuzzy values crisp rules are fuzzified. The expert system used these fuzzy values and generate fuzzy rules. The fuzzy output was then defuzzified to obtain a single output, which is a crisp value to indicate the stage of heart disease.

Durai et al. [17] used the fuzzy inference system in order to diagnose lung cancer. In this research domain specific dataset is prepared which includes features such as symptoms, levels of malignant cells and recommended treatments which helps in efficient diagnosis of the lung cancer. Soni et al. [18] proposed a weighted associative classifiers for prediction of heart disease. This research work was simulated using Java platform and MS access as a data server. This research work classified the patient records in two classes i.e. normal and abnormal case.

Neshat et al. [19] developed a diagnosis system using fuzzy logic for liver disorders. This paper proposed the binary classification system for classification of patient records into healthy or non-healthy patients. Kadhim et al. [20] implemented a back pain diagnosis system using a fuzzy expert system. In this research work, first of all different decision trees are created and fuzzy rules are designed for these decision trees in order to effectively classify the patient into healthy or non-healthy person.
Kalpana and Senthil [21] developed a diabetes diagnosis system using fuzzy rules. The simulation is performed using MATLAB. The experimental result is performed on the data of persons of age group of 26-30. Binary classification is performed in this work into healthy or non-healthy person.

R. Parvin and Abhari [22] presented the medical diagnosis system using FIS. This research follows the work in [15] using FIS to diagnosis disease. This research work is also based on the concepts presented in [16] for heart disease prediction.

Kalpana and Senthil [21] presented the research work for diabetes mellitus, and Neshat et al. [19] presented the research work for liver disorders. Both research work was presented using fuzzification rules to develop fuzzy inference system and simulated using MATLAB. Fuzzy rules use the features or symptoms in order to predict and detect diseases.

As stated in [16], [19] and [21], fuzzy inference system is used for feature detection or symptoms detection for heart disease. For example cholesterol level determines the heart disease level or type. The sensitivity of choosing variety of inputs was done by implementing the heart illness application in MATLAB once with a similar variety of inputs, output, and fuzzy rules employed in [16]. The result obtained from these research work shows the effectiveness of proposed work. All these experimental simulations are performed using MATLAB and MS Access for database. In the area of medical sciences, numerous expert systems have been developed. These are:

- PUFF: Pulmonary disease diagnosis
- VM: Monitoring of patients need to intensive care
- ABEL: Diagnosis of acidic materials and electrolytes
- AI/COAG: Blood disease diagnosis
- AI/RHEUM: Rheumatic disease diagnosis
- CADUCEUS: Internal medicine disease diagnosis
- ANNA: Monitoring and treatment analysis
- BLUEBOX: Depression diagnosis and treatment
- MYCIN: Microbial disease diagnosis and treatment
- ONCOCIN: Treatment and management of patient’s chemotherapy
• ATTENDING: Anesthesia management education
• GUIDON: Microbial disease education

1.5 Motivation for the Research

In these days, monitoring of different medical parameters can be done easily and all these data can be stored for future analysis and diagnosis. But the huge amount of data needs a lot of storage space which is costly. And another aspect is that all data are not of importance, only a few relevant data are important which posses a big problem in disease diagnosis. To make the disease diagnosis system more effective, the data should be filtered, conditioned and clustered. For clustering the data different contemporary and soft computing based data clustering techniques are available. After the data of different parameters are clustered, an inference mechanism is required to correlate two different parameters to a single cause.

The existing systems have various drawbacks like some were used for a particular type of dataset, some needed dataset of good quality. Therefore there is a need of a system of good quality that considers all the parameters, uses the best technique and predicts the diseases with greater accuracy. Fuzzy logic and expert system are important and very promising techniques in medical environment as it incorporates the knowledge and experience of physician and based on that information the system will predict the diabetes. With the help of fuzzy rule-based system we can avoid cost of conducting the test for the disease diagnosis. With this research we must be find out the accuracy of our system with the existing database so that we reach closely to create a perfect system for medical diagnoses

1.6 Objectives of the Research

The proposed system will solve the problem by selecting a subset of useful feature from a set of features. The main objective of the present study is to develop a control system to enhance the efficiency to diagnose a disease related to human disease. This work focuses on two severe diseases i.e. heart disease and dengue. The developed fuzzy expert system can explore crisp and linguistic data with loosely defined boundary
conditions for decision-making. It is implemented in MATLAB for the mentioned contexts for the comparison and validation with the dataset used in UCI repository and obtained from SAIMS Hospital, Indore. The proposed fuzzy controller makes the machine to take intelligent decisions as similar to that of humans.

1.7 Thesis Organization

Chapter 2: In this chapter, we describe the literature survey of the existing research related to our topic starting from the beginning till the current year and summarize the chapter.

Chapter 3: In this chapter a brief report of proposed fuzzy logic based algorithm have been designed for the diagnosis of heart related disease. The proposed expert system is designed by the help of fuzzy logic toolbox and used mamdani interface system. In this work, 10 input variables and 679 fuzzy rules are used over Cleveland Clinic Foundation database (200 patients) [23]. For defuzzification, centroid technique is utilized. It is one of the most efficient methods to create an expert system and diagnosis the heart disease.

Chapter 4: This chapter presents fuzzy expert system to raise the diagnosis level of dengue fever and early detection of dengue in patient. Fuzzy expert system is one of the most traditional artificial intelligence techniques to diagnose any disease.

Chapter 5: In this chapter a detailed discussion is presented on the result analysis done for the two identified diseases through MATLAB simulation environment. It contains the overall snapshots of our implementation in MATLAB. It is shown that the proposed work is quite better than existing work in terms of accuracy.

Chapter 6: This chapter describes the conclusion and the future work involved in our proposed techniques.