Fuzzy Logic for Medical Diagnosis

Thesis submitted in partial fulfillment of the requirements for the award of degree of

Master of Engineering

in

Computer Science and Engineering

Submitted By

Neeru Lalka

(Roll No. 801332016)

Under the supervision of:

Dr. Sushma Jain

Assistant Professor, CSED

COMPUTER SCIENCE AND ENGINEERING DEPARTMENT

THAPAR UNIVERSITY

PATIALA – 147004

July 2015

CERTIFICATE

I hereby certify that the work which is being presented in the thesis entitled, ***"Fuzzy Logic for Medical Diagnosis"***, in partial fulfillment of the requirements for the award of degree of Master of Engineering in *Computer Science and Engineering* submitted in Computer Science and Engineering Department of Thapar University, Patiala, is an authentic record of my own work carried out under the supervision of *Dr. Sushma Jain* and refers other researches' work which are duly listed in the reference section.

The matter presented in the thesis has not been submitted for award of any other degree of this or any other University.

Signature:
Neeru Lalka
(Neeru Lalka)

This is to certify that the above statement made by the candidate is correct and true to the best of my knowledge.

(Dr. Sushma Jain)
Assistant
Professor, CSED

Countersigned by

(Dr. Deepak Garg)
Head
Computer Science and Engineering Department
Thapar University
Patiala

(Dr. S. S. Bhatia)
Dean (Academic Affairs)
Thapar University
Patiala

ACKNOWLEDGEMENT

No volume of words is enough to express my gratitude towards my guide **Dr. Sushma Jain**, Department of Computer Science & Engineering, Thapar University, Patiala, who has been very concerned and has aided for all the materials essentials for the preparation of this thesis report. She has helped me to explore this vast topic in an organized manner and provided me all the ideas on how to work towards a research-oriented venture.

I am also thankful to **Dr. S. S. Bhatia**, Dean of Academic Affairs, **Dr. Deepak Garg**, Head of Computer Science & Engineering Department and **Dr. Ashutosh Mishra**, P.G. Coordinator, for the motivation and inspiration that triggered me for the thesis work.

I would also like to thank the staff members and my colleagues who were always there at the need of the hour and provided with all the help and facilities, which I required, for the completion of my thesis work.

Most importantly, I would like to thank my parents and the almighty for showing me the right direction out of the blue, to help me stay calm in the oddest of the times and keep moving even at times when there was no hope.

ABSTRACT

Medical diagnosis is a complex process due to the complexities, uncertainties and vagueness of the symptoms involved, and sometimes because of their indirect relationship with the final output. Traditional systems for diagnosis very often incorporate certain inabilities that eventually lead to the vagueness in the diagnosis result. Besides this, imprecise and incomplete knowledge are difficult for these traditional disease diagnosis expert systems to analyze. The fuzzy logic has carved a niche in medical diagnosis, for its ability to handle the dynamic nature of the disease diagnosis and medication. Various approaches of Fuzzy Logic, namely, Type-1 Fuzzy Logic, Interval Type-2 Fuzzy Logic, and General Type-2 Fuzzy Logic are being used for decision making in medical diagnosis. The fuzzy rule base is what that makes these approaches stronger. Various expert systems using these methodologies have been designed, however, an extensive study about how these different fuzzy based approaches serve the medical diagnosis that very often an evolvement of the basic Type-1 FL is demanded, leading to the formulation of 'layered Type-1 FL' approach called Interval Type-2 Fuzzy Logic, needs to be conducted to mathematically analyze the difference in the functioning of these approaches. In the thesis, these two approaches, i.e., Type-1 Fuzzy Logic and Interval Type-2 Fuzzy Logic, are implemented on disease dataset and are compared in terms of the accuracies of their predictions for two prominent lifestyle diseases, namely, diabetes and heart related complications. Type-1 Fuzzy Logic performs fuzzification using trapezoidal membership function, then rule inference and rule aggregation using MAX fuzzy based disjunction method is performed, and then the defuzzification using the centroid method. Interval Type-2 Fuzzy Logic performs type-2 fuzzification using the trapezoidal membership function with a uniform FOU, then rule inference is performed using MAX fuzzy based disjunction method. After that type-reduction using Karnik-Mendel algorithm computes the type-1 rule aggregation output, and finally the defuzzification using the centroid method computes the final output, i.e. probability. In this way, a disease specific performance of these fuzzy based methodologies is studied which is helpful to understand their usage and the functioning distinctively.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

Bmi:	Body mass index
C:	Constant of uncertainty
Cp:	Chest Pain Type
Df:	Diabetes function
D:	Diabetes
Exang:	Exercise induced anigma
Fgt:	Fasting glucose test
FL:	Fuzzy Logic
FoU:	Footprint of Uncertainty
G:	Gender
Mbp:	Minimum blood pressure
Num:	Angiographic disease status
Oldpeak:	ST depression induced by exercise relative to the rest
P:	Probability
Sft:	Skin folds thickness
Sit:	Skin folds thickness
Slope:	Slope of the peak exercise
T1FL:	Type-1 Fuzzy Logic
T2FL:	Type-2 Fuzzy Logic
Thalach:	Maximum heart rate achieved
Tid:	Total insulin dosage
Trestbps:	Resting blood pressure

LIST OF TABLES

1. INTRODUCTION

1.1 Medical diagnosis

Medical diagnosis and medication prediction, both are prone to various kinds of errors. Indirect and unknown relationships of the symptoms with the final output, is very difficult to recognize and work with. Experts and doctors often diagnose the disease inaccurately when the symptoms overlap with that of some other disease. Besides this, temporal nature of the symptoms may go unnoticed by the traditional diagnostic systems that basically operate on discrete information. The time-led variation in a disease, changes its stage and demands up-to-date medication. In the era of telemedicine and ubiquitous self diagnostic systems, various anomalies in the prediction may go unnoticed in the absence of an expert or doctor.

The problem also lies with the dynamism of the treatment. With the changing stage of a disease, treatment needs to be updated which may give better results to the patients. Maintaining good accuracy becomes difficult when a diagnosis involves a large dataset and vague knowledge. Figure 1.1 shows some of the complexities involved in the medical diagnosis.

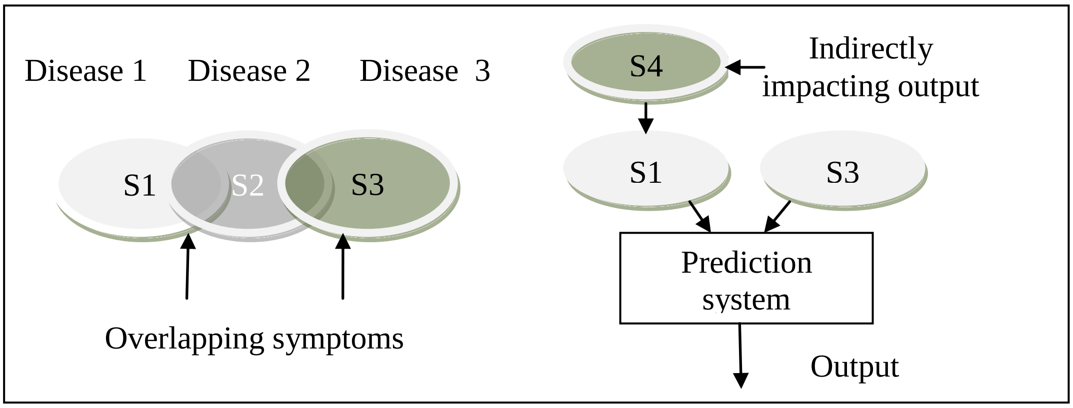

Figure 1.1 Uncertainty and complexity of symptom-disease relationship

Fuzzy logic has the tendency to handle the vagueness and uncertainty in a data set of any size. At the present scenario, importance of fuzzy logic based intelligent systems is recognized very well as it provides better accuracy and adapt to changes in the patterns of the disease diagnosis.

Lifestyle diseases are the long-term health conditions and require the medication to be administered over a long period of time. In such situations, it is imperative to accommodate the slightest change in the symptoms. The type-2 diabetes is attributed to the limited secretion of insulin and other factors like hereditary characteristics, diet, occupation, lack of workout, etc. are responsible for making diabetes condition worse. Some quantitative factors like glucose concentration in the body, body mass index, weight, etc. also affect the state of diabetes in an individual. Various intelligent approaches, namely, neural networks, fuzzy logic and genetic algorithm are being used to efficiently tackle the complexity of the attributes and for the optimizing rule base.

In the heart diseases, the problem lies with the uncertainty of various risk factors, namely, resting blood pressure (trestbps), maximum heart rate achieved on arrival at the hospital (thalach), level of depression induced by exercise as compared to the value when at rest, i.e. oldpeak, etc. What encourages the use of fuzzy logic for diagnosis of heart related complications, is the vagueness of distinction between a healthy individual and an unhealthy individual. Type-1 FL has been extensively used for diagnosing the heart related problems, whereas there is a limited use of the Interval Type-2 FL. The thesis focuses on the use of Interval Type-2 FL for predicting heart diseases.

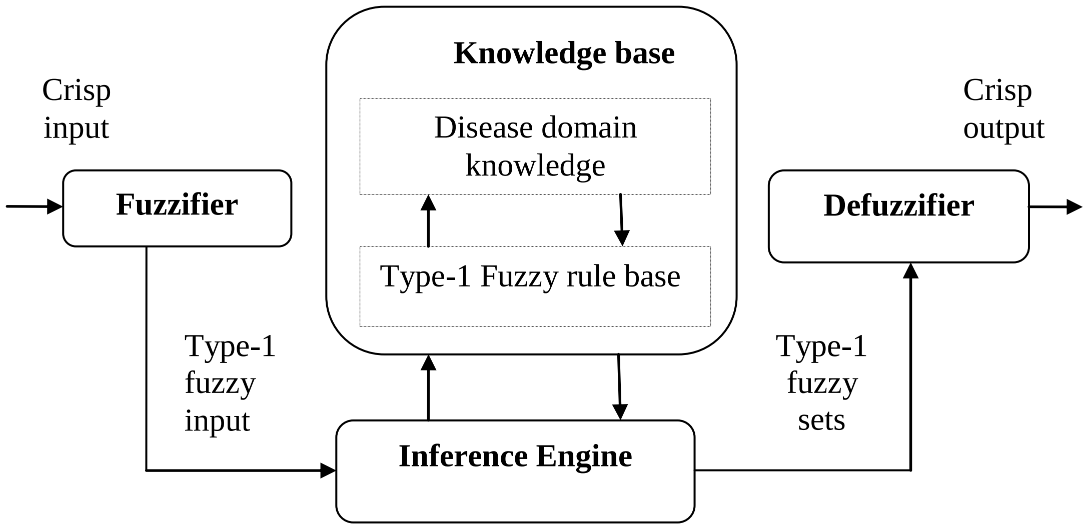

Figure 1.2 Type-1 Fuzzy Logic based expert system for disease diagnosis

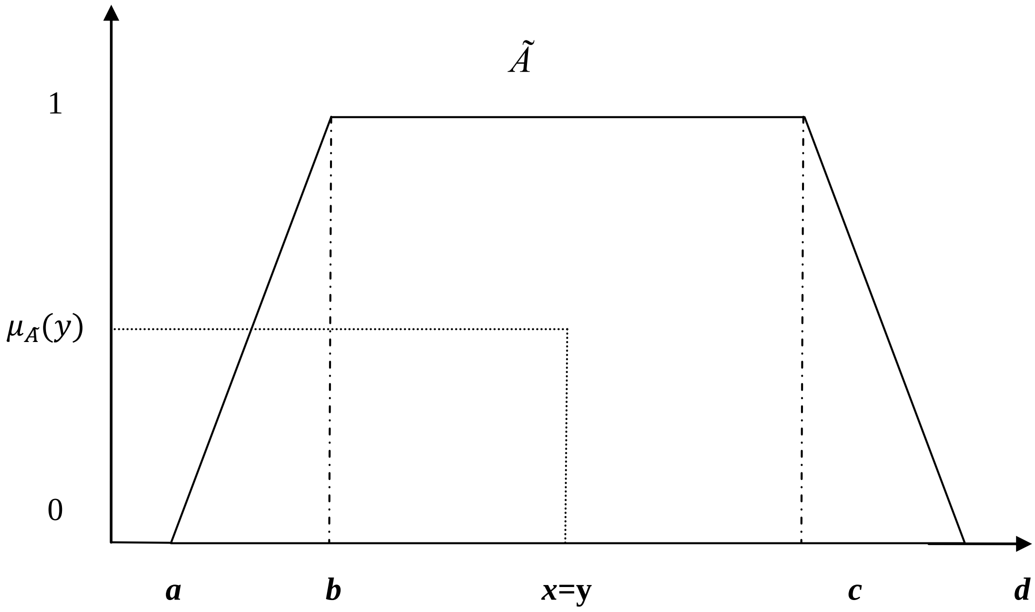

Figure 1.3 Trapezoidal membership function applied to crisp set x

1.2. Fuzzy logic

Fuzzy means uncertain information, and the mathematical logic that is used to handle that fuzzy data, is known as Fuzzy Logic (FL). Fuzzy logic means 'computation with the fuzzy sets'.

1.2.1 Type-1 Fuzzy Logic

Type-1 FL is the basic fuzzy logic methodology that performs fuzzification using the membership functions of various types depending on the requirement of an application. Fuzzy rule base basically stores the attributes as their fuzzy values where each fuzzy value incorporates a range of fuzzy values. The components of T1FL are shown in figure 1.2.

I. Fuzzifier: Fuzzification changes the crisp values to their corresponding fuzzy values using a membership function. A membership function incorporates the mathematical model and receives the crisp input to generate the corresponding fuzzy values which in turn belong to various fuzzy numbers. For a given problem, the type of membership function to be used depends on the number of fuzzy numbers. If we have two fuzzy numbers, the triangular membership function can be used and if we have three fuzzy numbers, we can use trapezoidal membership function as given in figure 1.3. The output, i.e. $\mu_A(y)$ is the fuzzy value for a particular fuzzy number.

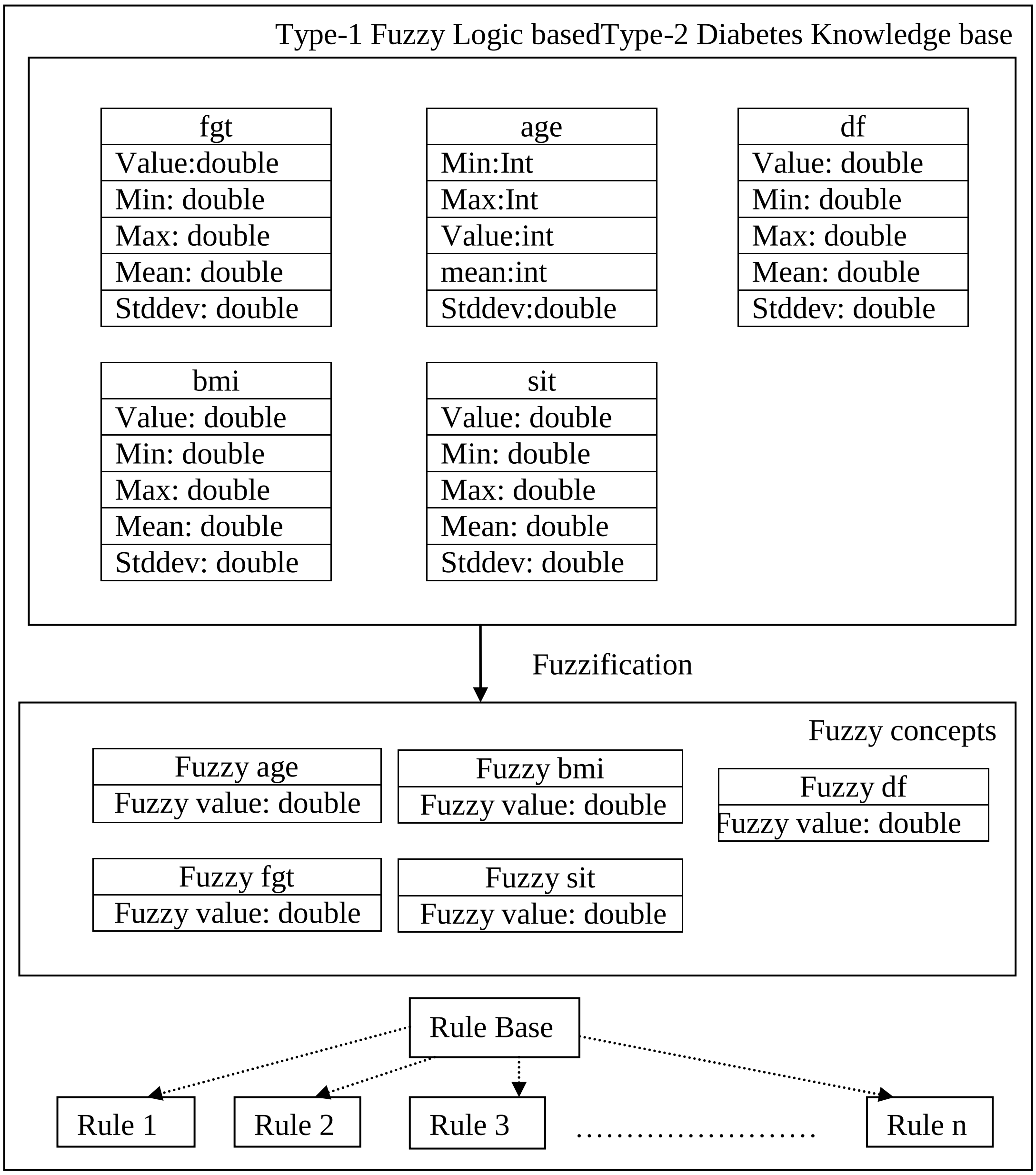

Figure 1.4 Knowledge base with Type-2 Diabetes as domain

II. Knowledge base: The knowledge base constitutes the domain specific knowledge. For instance, in fuzzy based Type-2 diabetes diagnosis, and insulin dosage control system, the domain is Type-2 diabetes. Knowledge base, as shown in figure 1.2, constitutes the rule base and the domain knowledge. Domain knowledge has both crisp and fuzzy values of the attributes, as shown in figure 1.4, and the rule base constitutes the type-1 fuzzy rules as if-then-else statements. Rules may have a particular class variable like diabetes mellitus and total insulin dosage for diabetes diagnosis and medication, respectively, and num in case of heart related diagnosis.

III. **Fuzzy Inference Engine**: Fuzzy rules can be formulated by data mining algorithms like apriori algorithm or a J48 algorithm. These rules are stored in the knowledge base, from where it is imported into the inference engine when a valid input is received. Here, the matching degrees are calculated for each fired rule using AND fuzzy based conjunction operation. Then fuzzy rule aggregation is performed by using MAX fuzzy based disjunction operation for two consequences, i.e. for positive diagnosis (1) and negative diagnosis (0).

IV. **Defuzzification:** Once the aggregated inference fuzzy values are computed, they need to be converted into their corresponding crisp values, i.e. the probability (0-1). There are numerous methods for performing defuzzification, like the mean of the max method, center of sum method, and the centroid method. These methods calculate the middle value in the overlapped area of different attributes after aggregation. Let D be the result of defuzzification after we apply the centroid method as given in (1). This value is the final output of the type-1 fuzzy logic, as shown in figure 1.5.

$$D = \left(\sum_{i=1}^{9} \mu_{\tilde{A}(xi)} * xi\right) / \left(s \sum_{i=1}^{9} \mu_{\tilde{A}(xi)}\right) \qquad (1)$$

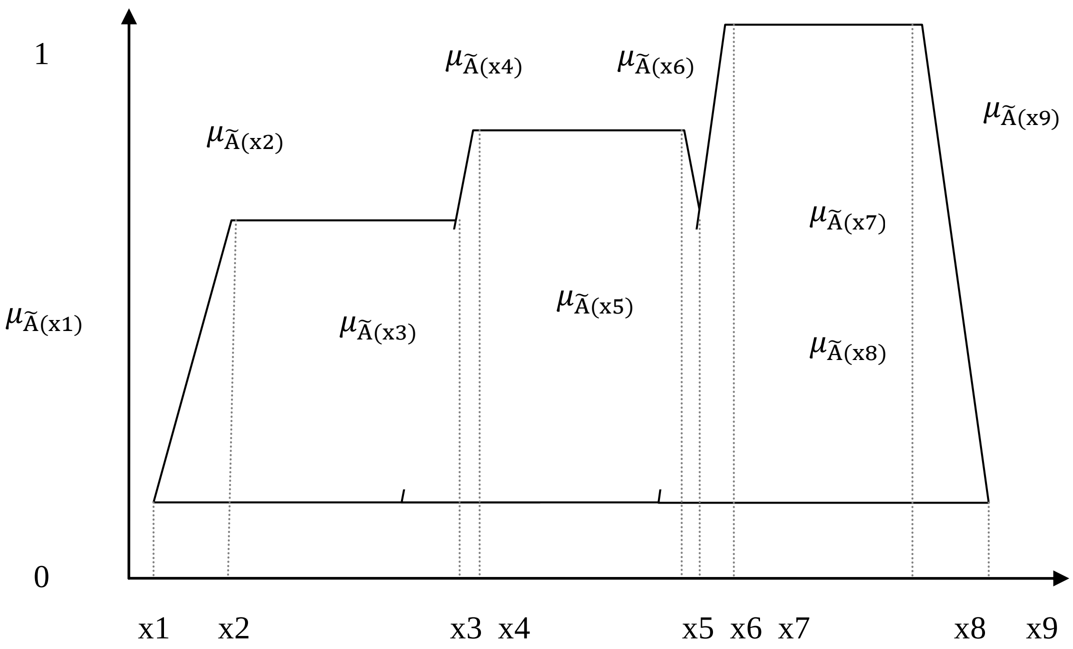

Figure 1.5 Defuzzification with 3 attributes using the centroid method

1.2.2 Interval Type-2 Fuzzy Logic

Interval Type-2 FL is a layered architecture of Type-1 FL. In Interval Type-2 FL, we have a fuzzifier that fuzzifies the crisp value into its type-2 fuzzy variant i.e. upper fuzzy value and lower fuzzy value, plus the corresponding type-1 fuzzy value. Rule inference is performed on these three fuzzy values for each consequence, i.e. 0 and 1. Type-reduction is performed to change these type-2 fuzzy values into their corresponding type-1 fuzzy values. A brief description is shown in figure 1.6.

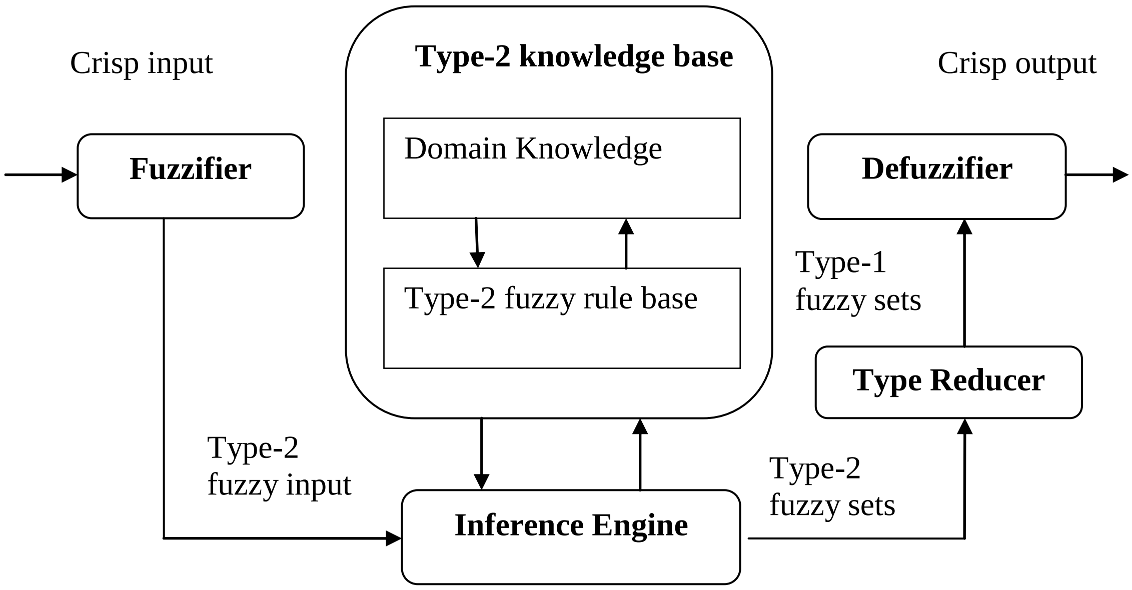

Figure 1.6 Description of Type-2 Fuzzy Logic System

I. Interval Type-2 fuzzifier: In Type-2 FL, instead of a single fuzzy value, we deal with three fuzzy values, namely, the upper fuzzy value of an attribute ($\bar{\mu}_{\tilde{A}}$), the lower fuzzy value ($\underline{\mu}_{\tilde{A}}$) and the type-1 fuzzy value of the same attribute ($\mu_{\tilde{A}}$). These three values are obtained from the Footprint of Uncertainty (FOU) around the trapezoidal membership function, as shown in figure 1.7. An uncertainty constant C is proposed to convert the irregular FOU to the regular FOU.

II. Type-2 fuzzy knowledge base: The knowledge base consists of the domain of type-2 fuzzy disease dataset and type-2 fuzzy rule base. The only change that is encountered in type-2 knowledge base is storing the fuzzy values of upper and lower trapezoidal membership functions. So, in interval type-2 FL the knowledge base expands.

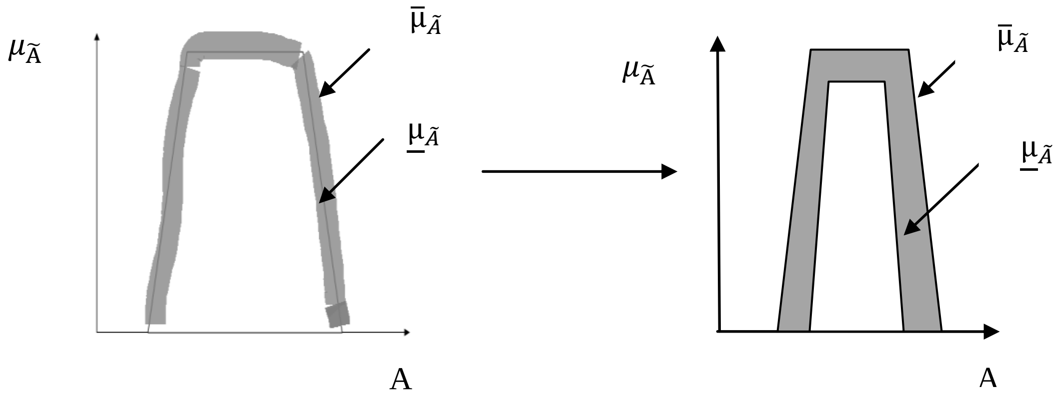

Figure 1.7 From irregular FOU to a trapezoidal membership function with uniform FOU

III. **Interval Type-2 FL inference engine**: The result of the fuzzification of the input values provides us with three fuzzy sets, i.e. $[\mu_{\widetilde{a1}}, \mu_{\widetilde{a2}}, \ldots, \mu_{\widetilde{an}}]$, $[\bar{\mu}_{\widetilde{a1}}, \bar{\mu}_{\widetilde{a2}}, \ldots, \bar{\mu}_{\widetilde{an}}]$ and $[\underline{\mu}_{\widetilde{a1}}, \underline{\mu}_{\widetilde{a2}}, \ldots, \underline{\mu}_{\widetilde{an}}]$. The type-2 fuzzy based rules are imported into the inference engine. Then rule inference is performed by using MIN fuzzy based disjunction operation for two consequences, i.e. for a positive diagnosis, ‘1’ and for negative diagnosis, ‘0’, in the disease diagnosis.

IV. **Type-Reduction**: The type-reduction is performed to convert the type-2 fuzzy based rule inference value to its type-1 fuzzy based variant. The type-reduction involves calculation of two switches, i.e. left switch ($\bar{L}$) and the right switch ($\underline{L}$) for each attribute in the FOU. The values of these switches are computed using the Karnik-Mendel type-reduction algorithm. In the end, as a result, we have the type-1 fuzzy sets for the different consequences.

V. **Defuzzification**: This step deploys the centroid defuzzification method to find out the probability.

1.2.3 Type-1 FL and Interval Type-2 FL for dynamic medication

The thesis discusses the difference of working of Type-1 FL and Interval Type-2 FL in predicting the medication required for diabetes diagnosis. The dynamism of the medication pattern shows the strength of FL over the traditional dosage recommendation system. The insulin dosage (units/kg/day) is recommended by taking into account the two attributes, i.e., plasma glucose level and body mass index (bmi). Further, using the simulation results, the strengths and weaknesses of Type-1 FL and Type-2 FL are identified.

1.3 Structure of the thesis

Chapter 2 discusses various types of works related to the fuzzy logic in medical diagnosis. Various approaches like Type-1 FL, Type-2 FL, data mining, ANN, GA, PSO, etc. are discussed which are used in the medical expert systems. Chapter 3 discusses the importance of the proposed work and its comparison with the previous works in fuzzy logic for medical diagnosis. Chapter 4 presents the mathematical model and steps of implementation. Chapter 5 discusses the analysis of the results obtained from the expert systems and compares type-1 FL and interval type-2 FL on various metrics. Chapter 6 presents the conclusion and the future scope of the proposed work.

2. STATE OF THE ART

2.1 Expert systems for medical diagnosis

Yau and Sattar [1] have demonstrated the use of soft computing techniques for the development of expert systems. The user interaction module acts as a prominent factor in decision-making. The author discussed the need of dealing with the complexity and dynamism of the knowledge to accommodate the changing patterns of the user requirements. Conigliaro *et al.* [2] have proposed an expert system for diagnosing venous insufficiency in humans using statistical evaluation to detect a set of the most influencing symptoms from a large data set. Levi [3] has proposed that it is imperative for the expert systems to be more accurate than human experts. He proposed some evaluation methodologies to figure out whether an expert system is beneficial in terms of the accuracy of the output. The two properties of these systems are analyzed by the author: its ability to predict, and the ability to complete the incomplete information. An expert system can also put its negative features in the final output, and hence a proper evaluation procedure is required to be set up.

Ahmed *et al.* [4] have proposed a study of the expert systems, in terms of their key features and shortcomings, and suggested guidelines for the development of new expert systems. The authors have compared expert systems based on the domains and language used for coding. Rahaman [5] has proposed a diabetes diagnosis system using JAVA Netbeans 7.1 Graphical User Interface (GUI) with SQL server for database management. The system covers all aspects of telemedicine like asking the user for the presence of risk factors, the results of the medical evaluation, symptoms and other questions about lifestyle and diet. Kim and Bekey [6] have proposed a methodology for applying the control scheme to a medical expert system to manage the abstraction of the database. The control procedure autonomously shifts from one level of abstraction to another, depending on the changing requirements for the output. At the same time, we have to determine the level of granularity of knowledge base, and inclusion property and common sense knowledge.

2.2 Soft computing methodologies for medical diagnosis

Different methodologies like Neural Network (NN), Fuzzy Logic (FL), Genetic Algorithm (GA), Ant Colony Optimization (ACO), and Particle Swarm Optimization (PSO) have been extensively used in medical diagnostic expert systems, specifically in order to mitigate complex mathematical calculations and for the optimization of final outputs.

2.2.1 Neural Network

Lakshmi and Padamavathamma [7] have proposed a methodology to identify gestational diabetes mellitus using feed-forward neural network. The method supports the self-examination. It evaluates the output using back-propagation neural network methodology. Temurtas *et al.* [8] have proposed the use of multilayer neural networks and probabilistic neural networks to solve the problem of classification of diabetes. Multilayer neural networks have replaced conventional pattern classifiers. The probabilistic neural network is based on "winner takes all" learning algorithm and uses the probability density function. Dey *et al.* [9] have proposed an artificial neural network approach with back-propagation technique for diabetes diagnosis. Multiple epochs can be used to reach the least possible error. As we increase the number of epochs, the error decreases. And also by increasing hidden layers and the number of neurons, less number of epochs are used to reach least error.

Aibinu *et al.* [10] have demonstrated the integration of Complex-Valued Neural Network (CVNN) approach with Real-Valued Neural Network (RVNN). The CVNN learns by experience and enhances the use of RVNN. Both approaches classify the input and helps in developing the input partitioning system.

2.2.2 Type-1 Fuzzy Logic

Fuzzy logic can be used in various ways in medical diagnosis, sometimes alone and sometimes, in collaboration with other soft computing methodologies. The existing work in the fuzzy based medical diagnosis incorporates:

i. Fuzzy ontology for developing expert system
ii. Data interpretation using fuzzy logic
iii. Fuzzy Logic for temporal diagnosis
iv. Fuzzy medical diagnostic systems
v. Neuro-fuzzy medical diagnostic systems

Lee and Wang [11] have proposed diabetes fuzzy ontology for diabetes diagnosis. The ontology first identifies the class membership of the patient as per his/her diagnosis inputs, and after that an individual specific diagnosis is practiced. The knowledge base is created using concepts, and fuzzy relations using fuzzy variables and fuzzy numbers. Calegari and Sanchez [12] have proposed the fuzzy ontology layered architecture. Fuzzy ontology uses semantic correlation to define fuzzy relationships, and defines knowledge base efficiently.

Abadi *et al.* [13] have proposed the use of recursive least squares in order to find the relationship between inputs and the outputs in a diabetes expert system. This kind of depiction of fuzzy relationships is flexible and the model can be evolved as many times as needed. Also uncertainties related to various aspects of the weight, time, quantity of meals, lifestyle can be considered while recommending diet. Gadoras and Mikhailov [14] have proposed input categorization mechanism for partitioning input. Input categorization involves generating rules, verifying them using a verification framework; knowledge and class representation using fuzzy sets, and rules help to identify the related classes in the case of identification of the new symptoms. The authors proposed the use of fuzzy logic for handling uncertainty in the class memberships.

Yager and Petry [15] have proposed a theoretical and logical approach for data summarization. Summarization enables the use of the concept /attribute hierarchies to control the information that is to be used in decision making. Fuzzy set approach is used to construct the concept/attribute hierarchies. Summarization should satisfy various parameters like minimum coverage, minimum relevance, succinctness, and usefulness related to a dataset. Levels of the summarization can also be fuzzy in nature.

Palma *et al.* [16] have proposed an approach to evaluate a disease in time-specific manner. The procedure considers the dynamic features of diagnosis prediction, their features and the patterns of dynamism. Then the parameters are modelled by fuzzy based Temporal Constraint Networks (TCN). Seising [17] has proposed the use of fuzzy sets and fuzzy relations for representing vague classes efficiently. A better representation of knowledge leads to an effective decision making and hence, a more reliable and efficient Computer Aided Diagnostic system can be developed. Straszecka [18] has proposed a fuzzy based probability approach to handle the uncertainty in diseases and the vagueness of the symptoms. The model represents symptoms as focal points, and uncertainty and significance of a particular symptom in terms of a value from 0 to 1, '0' being the least significant and '1' being highly significant. These focal points are represented by fuzzy sets. The probabilities of all the symptoms are added to measure the risk of the disease.

Kalpana and Kumar [19] have proposed fuzzy determination procedure for improving the accuracy and simplicity of diabetes diagnosis. Fuzzy determination mechanism (FDM) consists of fuzzy implication, fuzzy inference, and fuzzy aggregation. Das and Kar [20] have proposed the use of fuzzy soft approaches for designing diagnosis algorithms. The algorithm involves analyzing the considerations of a group of experts who give their opinion only regarding the symptoms they have knowledge about. Their opinion is assigned confidence using the soft fuzzy approach, and later on, these confidence values are aggregated to make an accurate decision. Mahfouf *et al.* [21] have proposed the fuzzy logic as an optimal solution to model the complex biological problems, medical diagnosis being one of them. Experiments on the use of fuzzy logic in biological problems started in 1980s.

Innocent and John [22] have proposed a computer based fuzzy diabetes diagnostic system. The expert system incorporates temporal information of a disease at various stages. Symptoms are studied against a constraint-satisfaction system. According to which only if all the symptoms collectively satisfy the positive diagnosis at a particular time, the patient is a diagnosed as diabetic (in case of diabetes). This system can be developed using type-2 fuzzy logic.

Kahramanli and Allahverdi [23] have proposed an approach using neural networks and fuzzy logic together. ANN reduces the complexity due to multiple dimensions of symptoms and is a cost effective approach. Using neuro-fuzzy model, continuous data can be represented and do classification efficiently. Gupta *et al.* [24] have proposed the use of fuzzy logic with neural network for the treatment of diabetes. Fuzzy rules are formulated and their fine tuning is done by NN. Neuro-fuzzy approach is used for learning and adaptation of diagnosis framework. A client-server approach is proposed for knowledge acquisition. Anouncia *et al.* [25] have proposed the use of rough sets for accurate knowledge representations. Use of fuzzy sets for knowledge representation leads to effective decision making, reduces the time and increases the accuracy.

2.2.3 Interval Type-2 Fuzzy Logic

There are certain fields where Type-2 fuzzy logic is being used prominently as given below:

i. Interval Type-2 FL based ontology
ii. Type-reduction algorithms

Lee *et al.* [26] have proposed a type-2 fuzzy ontology in which an ontology model is mapped on Type 2 Fuzzy Sets (T2FS). The T2FS is useful for personal diet recommendation. Type-1 Fuzzy Sets deal with crisp values and hence cannot deal with uncertainty and vagueness. It has Type-2 Fuzzy Personal Profile Ontology (T2FPPO), Type-2 Fuzzy Food Ontology (T2FFO), and Type-2 Fuzzy Profile Food Ontology (T2FPFO). The Intelligent Diet Recommendation System (IDRA) is based on T2FS. Mendel *et al.* [27] have proposed I-T2FL as more powerful than T1FL, in terms of dealing with fuzziness. Type-2 fuzzy is more accurate as they model uncertainties in membership functions too. The author also discussed General Type-2 Fuzzy Logic (GT2FL) which can be formed by layers of I-T2FL. The Karnik-Mendel (K-M) type reduction algorithm is used for non-fuzzy problems too.

Yeh *et al.* [28] have proposed a modified type reduction algorithm in place of the K-M algorithm. The author highlighted the inefficiency of initialization of switch points in K-M algorithms. The new algorithms, iteratively computes switch points and speeds up the process of convergence with previously calculated switch points. On a given universal set, $A = \{A_1, A_2, A_3, ... A_n\}$ and fuzzy set X, where $A_1 < A_2 < A_3 \ldots\ldots < A_n$, and $\mu_x(A_1)$, $\mu_x(A_2)$, ……..$\mu_x(A_n)$ are the intervals for $[\underline{I}_1, \overline{I}_1]$, $[\underline{I}_2, \overline{I}_2] \ldots [\underline{I}_n, \overline{I}_n]$ respectively, $\underline{I}_1$ being the lower membership function and $\overline{I}_1$ as the upper membership function for A_1 and so on.

According to Liu's algorithm for type reduction, the left switch $\underline{L}$ and right switch $\overline{L}$ values should be as given in (2).

$$\overline{L}_i = \overline{L}_{i+1}, \underline{L}_i = \underline{L}_{i+1} \tag{2}$$

This is how Liu's algorithm avoids regular iteration done by the K-M algorithm to compute switch points. Aladi [29] has proposed the efficient calculation of the FoU for the triangular membership function to provide a uniform range of uncertainty over the whole membership function. Using the Uncertainty Indicator, one can efficiently calculate the left and right switch points.

2.2.4 Genetic algorithms

Roychowdhury *et al.* [30] have proposed an expert system using fuzzy logic controller and genetic algorithm (GA). The fuzzy controller forms the rule base in online mode. The rule-base is subjected to genetic algorithm for fine tuning and for optimizing membership functions in the knowledge base. Optimized rules are verified by the experts. GA performs the selection process on a rule base in the model. Ling *et al.* [31] have proposed genetic algorithm based fuzzy multiple regression technique for detecting hypoglycemic episodes. GA selects optimized parameters of the regression, and fuzzy inference system identifies the episodes. Chan *et al.* [32] have proposed a method for classification of hypoglycemic episodes using hybrid of neural networks and genetic algorithm. NN and GA develop an efficient classification model for detection of hypoglycemic episodes. But due to the black box nature of NN, experts feel uncomfortable with the logic of output formulation by neural networks. In the proposed model, the outputs are validated by experts.

2.2.5 Particle Swarm Optimization

Hsieh *et al.* [33] have proposed Particle Swarm Optimization (PSO) for pruning the rules fired by the fuzzy inference system. NN is difficult to interpret and hence, neuro-fuzzy rule base is difficult to evolve over a period of time. Both NN and FL suffer from the problem related to knowledge acquisition. Inbarania *et al.* [34] have demonstrated a feature selection approach based on PSO. PSO resists irrelevancy and redundancy in symptoms. Fuzzy set reduces the dimensions of a dataset of symptoms and enables accurate classification. The methodology involves a PSO approach to be used after fuzzy set formation. Kumar *et al.* [35] have proposed neural network optimized by ACO, which means an integration of ACO and feature selection. A comparative study of ACO and ACO with feature selection is presented in which NN optimized by ACO exhibits better accuracy and reliability in healthcare applications. Ganji and Abadeh [36] have demonstrated the use of ACO for fuzzy rule extractions to perform classification. ACO is an effective approach for data mining. The methodology also updates in pheromone rule for better decision making. And also the author emphasizes the cooperation rather than competition aspect of ACO.

2.3 Data mining in fuzzy expert systems

Kumari and Singh [37] have proposed the telemedicine for conducting diabetes diagnosis using data mining approach. The model uses neural networks for detection of diabetes mellitus (DM). The estimation of diabetes mellitus depends on the initial symptoms and some physical conditions. Patil *et al.* [38] have proposed the necessity of emphasizing on processing procedure in the initial phases of knowledge discovery in data mining approach. A modified equal interval binning approach is proposed for discretization of continuous values. Experts get to decide the binning width and the width is also given to the model as input. Discretization assigns categories to each value and an apriori algorithm is used to generate rules.

Kwiatkowska and Kielan [39] have proposed two methods for interpreting, designing and modeling the diverse medical database, namely, fuzzy logic for modeling continuous data, and semiotic methods to model symptoms based on their dimensionality and the specific context, for exploiting the dataset in computerized manner for a diabetes diagnosis. Complexity arises during modeling, interpretation and representation of medical parameters/concepts. Suh and Vudumula *et al.* [40] have proposed concept and attribute hierarchies with clustering algorithm for data mining in diagnosis of diabetes. The proposed approach is useful for early detection of diabetes in patients. However, if the data set size increases, the complexity of the system increases, making additional steps imperative for controlling complexity. Esfandiari *et al.* [41] have demonstrated a study on medical data mining. Medical data mining is important for knowledge acquisition in expert systems and improves the overall performance and accuracy The author categorizes various studies on medical data mining on the basis of different approaches like classification, hybrid technology, association and clustering.

2.4 Fuzzy Logic in medication

Delgado *et al.* [42] have proposed a closed loop approach for glucose-insulin control regime. Monitoring of insulin dosage is done on a daily basis and besides this a short duration dose control regime is followed before each meal in a day. Frequent measurement of glucose is done in order to achieve accuracy. A target of glucose level is set for each day, which can change according to variations in metabolism in an individual. Grant [43] has proposed using fuzzy as the artificial pancreas to regulate insulin dosage in a diabetic. The framework uses a feedback model with closed loop to regulate glucose in a diabetic with the help of biomedical sensors, instead of regular blood tests and other medical tests. The author used fuzzy to adjust the meager differences of the two glucose readings to administer insulin effectively.

Nazari *et al.* [44] have proposed a fuzzy model for controlling diabetes mellitus using a recursive least square method. The model involves controlling diet regime among the patients using inputs like weight, time of meals, duration of the simulation and glucose concentration in the fuzzy expert system. Body sugar and carbohydrate intake level is decided each day using MATLAB based glucoSim Simulator.

Adeli [45] have proposed the fuzzy based expert system for heart disease diagnosis. The author uses the Mamdami based inference method and compares the obtained results of the diagnosis with those in the database. The expert system designed here stimulates the way an expert and doctor relation to design a more accurate disease diagnosis system.

3. PROBLEM STATEMENT

3.1 Gap analysis between the previous work and the targeted work

The previous works discuss the use of fuzzy logic approaches in various diseases. Type-1 FL is used for: the medical diagnosis, data mining and managing the dataset in order to use summarized data. Further, it is also demonstrated that Type-1 FL can be used for controlling insulin dosage based on the varying glucose levels and for diet recommendation systems. Type-2 FL is mainly used for organizing database in the ontology as ontology deals with large and complex data set generated from different media. Besides that, in the field of medicine, Type-2 Fuzzy Logic is also used.

However, none of these approaches are ever used for medication prediction. There is a scope where Type-2 FL can be used for recommending medication. Also, the merits of Type-1 FL over Type-2 FL are hardly demonstrated and most of the times the type-1 fuzzy and type-2 fuzzy are being used interchangeably for a diagnostic system. Further, besides diabetes, there is no other disease where type-2 fuzzy logic has been used extensively. Due to the non-uniformity of the uncertainty curve around the membership functions, accurate output generation requires vigorous mathematical calculations.

The difference of the, previous works and proposed work, lies in the defining a distinctive nature of the functioning of Type-1 FL and Interval Type-2 FL. Looking at the architecture of the Interval Type-2 FL, it is expected to perform better than Type-1 FL for every disease diagnosis. But there are certain areas where Type-1 FL is more useful because of the easy implementation and speedy generation of results. When it comes to recommendation of insulin dosage, isn't it more useful if we use Interval Type-2 FL for calculation of probability of diagnosis followed by Type-1 FL for insulin dosage recommendation?

3.1.1 Interval Type-2 FL in complex disease diagnosis

In complicated diagnosis, one should use Interval Type- 2 FL, because of its layered architecture that makes it more accurate than Type-1 FL. Even then, it is observed that practically Type-1 Fuzzy Logic could perform better than interval Type-2 FL.

3.2 Type-1 Fuzzy Logic and Type-2 Fuzzy Logic on medical diagnosis

In the thesis, Type-1 FL and Interval Type-2 FL are used for predicting the diagnosis of diabetes and heart related complications. Various metrics have been used to compare the two approaches discussed as follows:

a) Diagnosis probability calculated by Type-1 FL and Interval Type-2 FL: For each disease the probabilities calculated by the two approaches are analyzed to find out the characteristics of the variations in their working. The difference in probabilities can range from 0 to 1-p1, where p1 is the probability calculated by T1FL, and 0 shows no difference between the probabilities of T1FL and I-T2FL and 1-p1 shows the great difference.

b) The nature of rule base: It is prominently argued that the more distinct the rule base is, the better is the approach. Here, we will compare the rule base formulated by the two approaches in terms of the total number of rules and the distinctiveness of the formulated rules.

c) Medication administration of insulin: It is unclear that whether to use T1FL or I-T2FL for insulin dosage recommendation. However, both the approaches are implemented to predict the dosage and are compared in terms of the level of accuracy.

3.3 Objectives of the proposed work

It has been proposed why Type-1 FL is still relevant despite the evolvement of Type-2 FL. As heart complications and diabetes are the two most complex diseases, so it looks imperative to use the Interval Type-2 FL for both. However, Type-1 FL can be also be used for these diseases due to its own advantages like easiness, fast output calculation.

a) To study if the difference in the outputs by these approaches is uniform or non-uniform with the increasing severity of the symptoms.

b) To understand the utility of Interval Type-2 FL for various medical problems where the use of Type-1 FL is widespread.

c) To formulate the collaboration of Type-1 FL and Interval Type-2 FL to make disease diagnosis more efficient. Each approach will help to overcome the weakness of the other.

4. PROPOSED WORK

4.1 Type-1 Fuzzy Logic based diabetes prediction

As stated earlier, diabetes diagnosis is performed using various interrelated symptoms/attributes. The knowledge base constitutes of Type-2 diabetes domain which undergoes fuzzification and the corresponding fuzzy values are also stored in the knowledge base. The PIDD database from American Diabetes Association is used for diabetes diagnosis.

Fuzzification is performed on the knowledge base using a trapezoidal membership function. The attributes are obtained from an individual and the final output about the probability is stored in D, the class variable.

I. Type-2 diabetes knowledge base: The knowledge base has domain specific knowledge and the corresponding fuzzy values. Besides this, it also has a fuzzy rule base. In the database before fuzzification, max, min, standard deviation and mean values are calculated for each attribute for all the instances and are then stored in the database. For e.g. if there are N number of instances, then *mean(age) = sum(age)/N.* Sum(age) is the sum of all the values corresponding to 'age' attribute. Then standard deviation is calculated for all the attributes and formula for standard deviation, for instance, age, is given in (3). From these results, [m,n,o,p] fuzzy numbers are calculated for age_low, age_med, age_high, fgt_low, fgt_med, fgt_high, sit_low, sit_med, sit_high, bmi_low, bmi_high, bmi_med, df_low, df_med, df_high, mbp_low, mbp_med and mbp_high where m is "begin", n is "after begin", o is "before end" and p is the "end".

$$stddev(age)=\sum_{i=1}^{N}\frac{\left(value(age)_i-mean(age)\right)^2}{N} \tag{3}$$

a) Fuzzification: In Type-1 FL, trapezoidal membership function is used as shown in fig. 1.3. The membership function for each attribute is computed for all the corresponding fuzzy numbers, i.e., low, med and high. Using (4), we will have three membership functions for each attribute, i.e., membership degree for low, med and high fuzzy numbers. This step is also performed in converting the crisp dataset into type-1 fuzzy based dataset.

$$x = \begin{cases} 0, x < m \ or \ x > p \\ \left(\frac{x-m}{n-m}\right), m \leq x < n \\ 1, n \leq x \leq o \\ \left(\frac{p-x}{p-o}\right), o < x \leq p \end{cases} \quad (4)$$

b) Type-1 fuzzy rule base: The rule base is formulated using the apriori algorithm of the WEKA data mining tool. The If-then-else rules are constructed using the five attributes, namely, fgt, sit, age, df and bmi. D which stands for diabetes mellitus, becomes a class or target variable and stores the result. The rules are constructed using a fuzzy domain knowledge base. One such rule is given below.

```
if (fgt=='high')
    if (bmi=='med' && df==''high')
        then D='tested_positive';
    else if (bmi=='med' && df=='low' && age=='low')
        then D='tested_negative';
```

II. Diabetes diagnosis decision making: The decision-making process starts when a valid input is received, i.e., I=[fgt, mbp, sit, sft, bmi, age, df]. The fuzzification is performed for input values, resulting in the construction of I'=[fuzzy fgt, fuzzy sit, fuzzy sft, fuzzy bmi, fuzzy age, fuzzy df, fuzzy mbp]. The decision making further involves rule matching, rule inference and rule aggregation, and defuzzification for generating the output.

$$r_{\mathrm{inf_i}} = MIN(fuzzy(age)_{i,} fuzzy(fgt)_{i\,,} fuzzy(sit)_{i,} fuzzy(bmi)_{i,} fuzzy(df)\ ) \quad (5)$$

a) Type-1 Fuzzification: The fuzzification of the input instance is performed as given in (4). The fuzzy values of all fuzzy numbers are stored in the dataset.

$$r_{agg} = \mathrm{MAX}\ (\ r_{inf}(i)) \quad (6)$$

$$\text{Dm} = \sum_{i=1}^{n}(w(i) * y(i)) / \sum \text{y(i)} \tag{7}$$

b) Rule Inference: A set of 60 rules has been used, so matching degree will be calculated for 60 times ranging from (0-1). After that the result of the inference of fired rules, that is the minimum of the matching degrees of the rules with similar consequences is calculated as shown in (5), where 1<=i<=N. The result is a set of rules for each consequence, i.e., $[r_{inf_i}, r_{inf_{i+1}}\ r_{inf_{i+2}} \dots \dots .. r_{inf_n}]$, where n is the number of fired rules. Then from these values, we calculate the rule aggregation value, i.e. max of the inferred rule values as shown in (9) for each consequence.

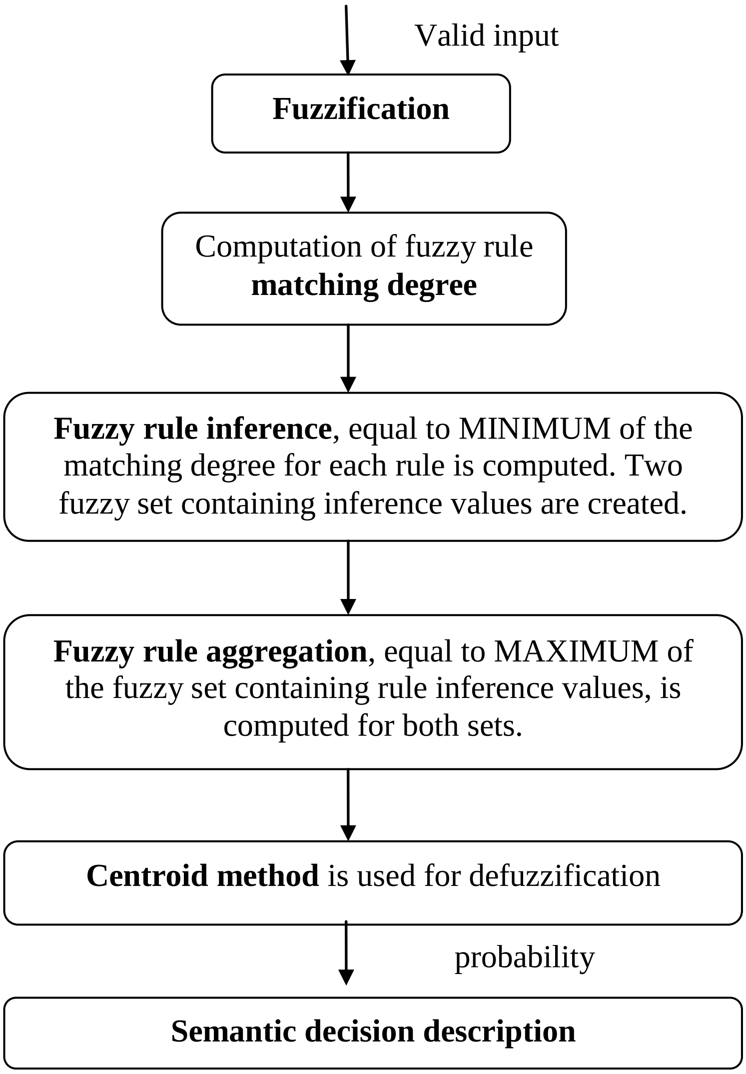

Figure 4.1 Flowchart depicting the steps of Type-1 FL medical diagnosis decision making

c) **Defuzzification:** Then using (7), i.e. centroid method, where *w (j)* is the weight of the aggregated rule value and j = [1, 2], the final probability is computed.

d) **Semantic description of the output**: The probability of diabetes is divided into five parts, i.e. very low, low, medium, high and very high; if the crisp value of probability lies between 0 and 0.20, the probability is very low, if between 0.20 to 0.40, it is low, if between 0.40 to 0.60, it is a medium, if between 0.60 to 0.80, the probability is high and if it is between 0.80 to 1 or equal to 1, it is very high.

4.2 Type-1 Fuzzy Logic for heart related disease prediction

Probability of heart related complications is calculated, where ‘0’ means negative diagnosis and ‘1’ means positive diagnosis. The steps of computing probability are similar as shown in figure 4.1. A heart related complications data set from the UCI machine learning repository is used as disease data set. Then Type-1 FL fuzzification is performed on the dataset, rule base is formed, and both the fuzzy dataset and the rule base are stored in the knowledge base.

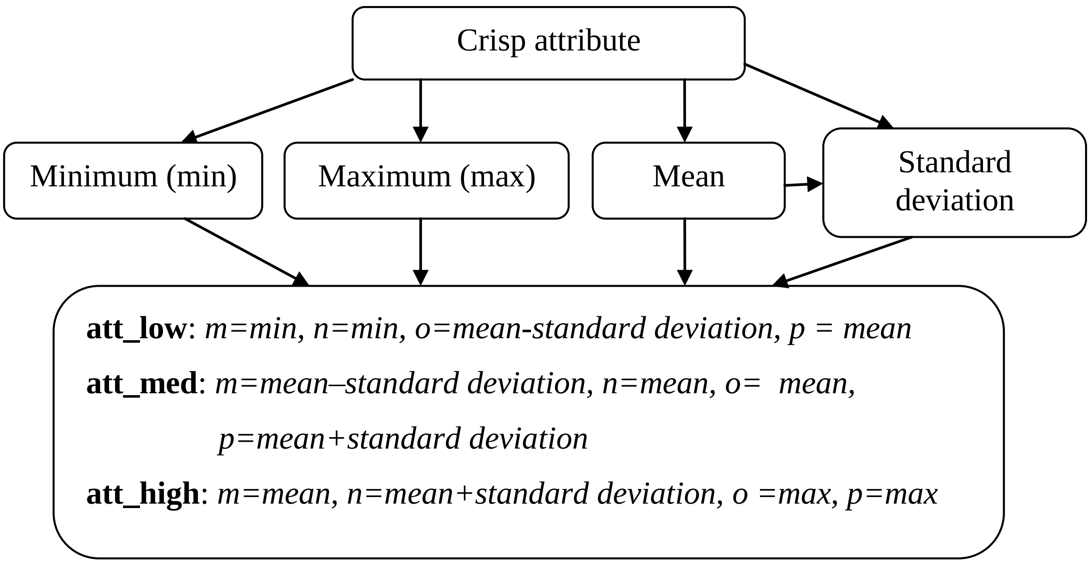

Figure 4.2 Construction of fuzzy numbers

I. **Fuzzy based knowledge base:** This dataset is fuzzified using the same procedure as used in type-2 diabetes diagnosis. Before fuzzification, the minimum, maximum, mean and standard deviation is calculated for all the attributes in the dataset. Then construction of low, med and high fuzzy numbers is done for all the attributes as shown in figure 4.2. After constructing fuzzy numbers, the crisp value is fed to the trapezoidal membership function for computing the corresponding fuzzy value for the fuzzy numbers. However, it is noticeable that for one of these three fuzzy numbers, the fuzzy value will be 0.

II. Rule base: A rule base is constructed using the WEKA data mining tool. Apriori algorithm of the association tool is used to construct 50 rules. The target class is the ‘num’ attribute which can have any value from the set [1,2,3,4]. Some of the rules so formed are shown below:

```
1. if (cp=='high’ && age='low' && trestbps='low' && thalach='med')
                  if (exang='low')
                         then num=2;
                  else if(exang='high')
                         then num=3;

2.  if (cp=='high' && age='low'&& trestbps='low' && thalach='low')
                  if (exang='low')
                         then num=4;
                  else if(exang='high')
                         then num=2;
```

III. Type-1 FL based prediction of heart related complications:

a) Fuzzification: The fuzzification is done using (4) when a valid input is received.

b) Rule matching degree: Input fuzzy values are tallied against the fuzzy rule base and rule matching degrees are calculated. The rules for which the matching degree is greater than 0 are the fired rules. There will be a different set of firing rules for the four different consequences, i.e., with respect to the four different values of ‘num’.

c) Rule inference, and aggregation: Rule inference is performed by using the MIN fuzzy conjunction or AND fuzzy conjunction. The minimum value of the attributes is chosen as the rule inference value. If there are n fired rules, then there will be n rule inference values. Rule aggregation is performed using MAX fuzzy based disjunction as shown in (6).

d) **Defuzzification:** Centroid method as shown in (7) is used for defuzzification. Here all rule aggregation values of the four consequences, i.e. num=1, num=2, num=3, num=4 are combined to get the final probability.

e) **Semantic decision making:** The probability of diabetes is divided into five parts, i.e. very low, low, medium, high and very high i.e., if crisp value of probability lies between 0 and 0.20, the probability is ‘very low’, if between 0.20 to 0.40, it is ‘low’, if between 0.40 to 0.60, it is ‘medium’, if between 0.60 to 0.80, the probability is ‘high’ and if it is between 0.80 to 1 or equal to 1, it is ‘very high’.

4.3 Interval Type-2 Fuzzy Logic based diabetes prediction

As shown in figure 1.6, in Type-2 FL we have three dimensional membership function and a type reducer to change the rule aggregation value to its type-1 variant. It works as a ‘layered Type-1 FL’.

I. **Construction of knowledge base:** The knowledge base consists of the domain knowledge of type-2 diabetes. As we deal with the upper and lower range of a simple type-1 fuzzy value, the knowledge base is expanded. For an instance of the diabetes dataset, we will have around 15 fuzzy numbers.

II. **Rule base construction:** Now, the rule base is constructed on this expanded FL based diabetes dataset. In this way, the rule base also expands, however, if the confidence percentage is considered the rule base of the Type-1 FL based dataset and Type-2 FL based dataset are almost same. Some of the rules of the Type-2 FL based rule base are shown below.

```
6.        if (df==’low’ && age== ‘high’)
              If (pgc==’high’ && bmi=’med’)
                  then D=’tested_positive’;

8.        if (df==’low’ && age== ‘high’)
              if (pgc==’med’ && bmi=’high’)
                  then D=’tested_negative’;
```

III. **Type-2 FL based diabetes decision making:** The steps are similar as type-1 fuzzy logic based decision making except the inclusion of type-reduction module before defuzzification. Now, as we discussed earlier, the crisp input is converted into its three dimensional variant.

a) Construction of FoU: FoU means Footprint of Uncertainty which exists around the boundaries of the trapezoidal membership function for all the attributes as shown in figure 1.6 and it can be uniform or non-uniform. In the proposed work we are using uniform FoU around the membership function. It is calculated using the upper and lower membership functions, which are shown in (8) and (9).

$$\overline{\mu}_{\tilde{A}} = min(\mu_{\tilde{A}} + \frac{C}{2}, 1.0) \tag{8}$$

$$\underline{\mu}_{\tilde{A}} = min\left(max\left(\mu_{\tilde{A}} - \frac{C}{2}, 0.0\right), 1.0 - c\right) \tag{9}$$

$$FoU = MAX\,(\overline{\mu}_{\tilde{A}}, \underline{\mu}_{\tilde{A}}) \tag{10}$$

(10) shows that FoU is the max of the two membership functions for a particular attribute. The value of the constant C which is used in the (8) and (9), should lie between 0 and 1 because the final probability lies between 0 and 1. C = 0 means negligible FoU and C = 1 means extremely wide FoU.

b) Calculation of Indicator of Uncertainty: Such an indicator helps to compute the level of uncertainty captured by the FoU for the attributes. It is usually the difference of the upper and lower membership function i.e. $\overline{\mu}_{\tilde{A}} - \underline{\mu}_{\tilde{A}}$.

c) Fuzzification: Trapezoidal membership function with FoU is used to find the fuzzy values of an instance as shown in (4). In this way, for each attribute we have three fuzzy values where the upper fuzzy value > T1FL fuzzy value > lower fuzzy value.

$$r_{\mathrm{inf_i}} = fuzzy(age)_i * fuzzy(sit)_i * fuzzy(bmi)_i * fuzzy(pgc)_i * fuzzy(df)_i \tag{11}$$

d) Rule inference for each consequence: The rule base is loaded in the fuzzy inference engine. The matching degree for each rule is calculated and only those rules are considered to be fired where the matching degree is greater than 0. After that, rule inference is performed by using t-form production method for each consequence, which means the production of the matching degrees as shown in (11). At the end of this step, we will have two sets of the fired rules, i.e. for 0 and 1.

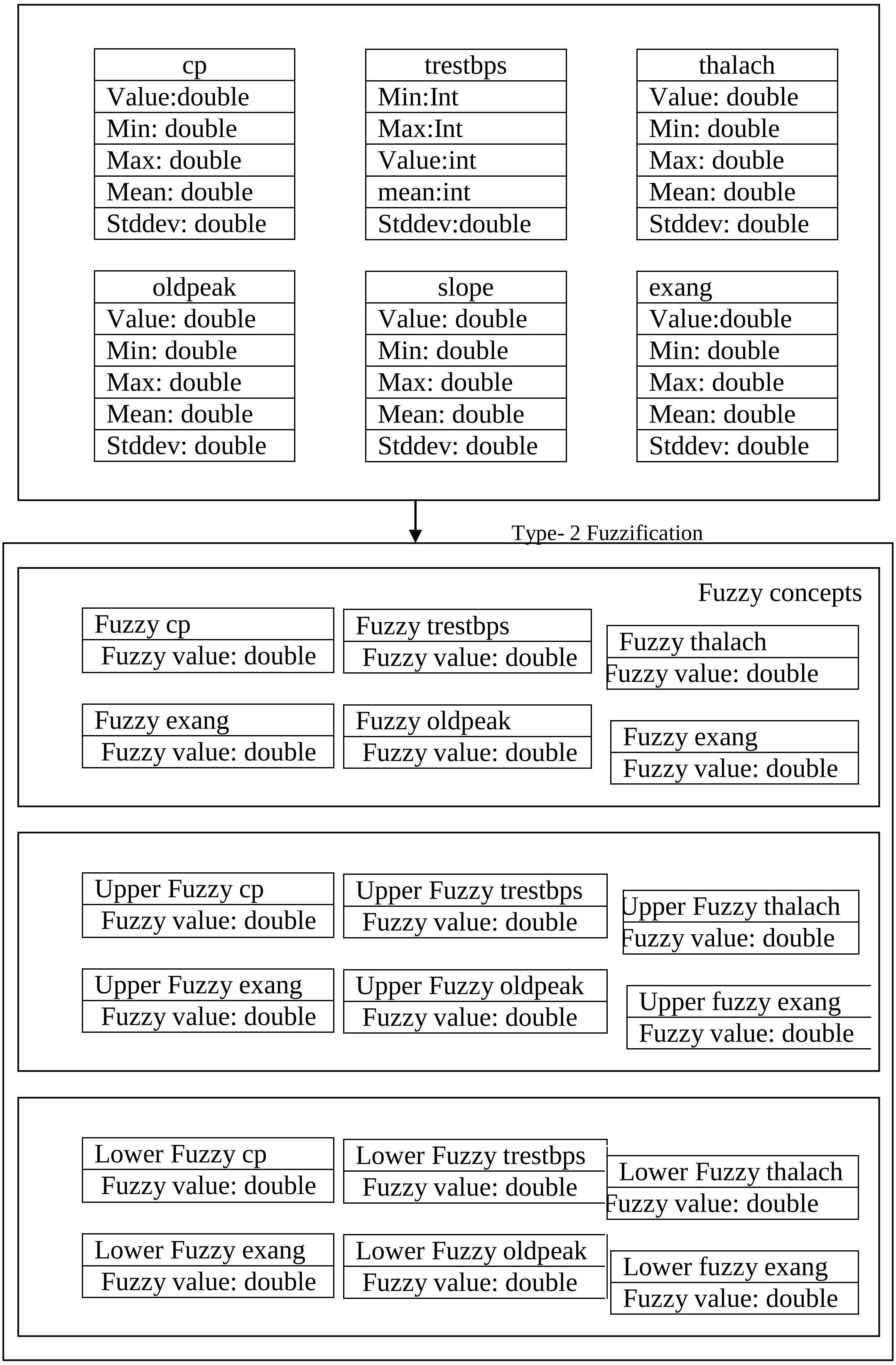

Figure 4.3 Interval Type-2 FL based heart related complications knowledge base

e) Type-reducer: The proposed work uses the Karnik and Mendel (K-M) type-reduction algorithm. The fired rules have inference values on which the type-reduction is performed. First step is the calculation of T for each consequence, i.e. switch points. $\underline{T}$ and $\overline{T}$ which are left and right switch points respectively.

$$P_j = (\overline{R}_j + \underline{R}_j)/2 \tag{12}$$

The initial values of the left switch points and the right switch point is given in (14) and (15), where n is the total number of fired rules. Then upper and lower type reduced fuzzy rule aggregation value is calculated as shown in (16) and (17). Here, $\overline{T}$ and $\underline{T}$ are the upper and lower type reduced values. 'i' points the particular rule on which type reduction is being performed. j points to the consequence for which type reduced upper and lower values are being computed and here j = [*0*,*1*]. x_i is the secondary membership function for each rule, i.e. type-1 fuzzy based rule inference value and $\overline{I}_{i,j}$ is the upper rule inference value, and $\underline{I}_{i,j}$ is the lower rule inference value. So, we will have two type reduced sets for diabetes diagnosis. Using (12), the Type-2 FL based rule aggregation values are computed for the two consequences and are are later defuzzified to get the probability.

f) Defuzzification: Using the centroid method as stated in (10), the final probability is calculated. Here, *w (i)* is the weight of aggregated fired rules for a consequence, *i* = [*0*, *1*]. And the *y (i)* is the corresponding type-reduced type-1 fuzzy value.

4.4 Interval Type-2 Fuzzy Logic for heart related complication prediction

Heart related complication diagnosis is performed just as a Type-2 FL based diabetes diagnosis. The difference lies in the domain knowledge. The proposed work uses a heart dataset from the UCI data mining repository. The brief discussion of the implementation of the Type-2 FL for heart related diagnosis predicted is written below.

I. Construction of knowledge base: Due to three dimensional fuzzification, the knowledge base expands as shown in figure 4.3. FoU for each attribute is calculated as the maximum of the upper and the lower membership function. The indicator of uncertainty is the difference of the two membership functions.

II. Type-2 FL based rule base for heart disease: The type-2 fuzzy rule base uses all the upper, lower and secondary membership functions to construct the optimal rule base. The Weka data mining tool is used for this. Rule no. 10 is shown below.

```
10.   if (age==’high’ && cp==’high’)
          if (thalach==’high’ && exang==’high’)
              then num=3;
          else if (trestbps==’high’ && thalach== ‘low’ && oldpeak==’med’)
            Then num=3;
```

III. Heart related disease prediction:

a) T2FL fuzzification: Trapezoidal membership function with FoU for each attribute is used to find the fuzzy values of an instance. In this way, for each attribute we have three fuzzy values, where the upper fuzzy value > T1FL fuzzy value > lower fuzzy value.

b) Rule inference for each consequence: The matching degree for each rule is calculated and only those rules are considered to be fired where the matching degree is greater than 0. After that, rule inference is performed using the t-form production method for each consequence, which means the production of the matching degrees as shown in (11). At the end of this step, we will have four sets of the fired rules, i.e. for num=0, num=1, num=2, num=3.

c) Type-reduction: The proposed work uses the Karnik and Mendel (KM) type-reduction algorithm. The fired rules have inference values on which the type-reduction is performed. The initial values of the left and right switch points are calculated as shown in (14) and (15). The upper and lower type reduced values are calculated using (16) and (17). Here j = [1, 2, 3, 4]. Using (12), the Type-1 FL based inference values are computed for the four consequences which are later defuzzified to get the probability.

d) Defuzzification: Using centroid method as stated in (7), the final probability is calculated.

4.5 Type-1 Fuzzy Logic based insulin dosage prescription

I. Type-1 FL based knowledge base: The knowledge base used for insulin prediction has two attributes i.e. bmi and pgc. The class variable is tid. The fuzzy numbers for insulin dosage are computed using the minimum, maximum, mean and standard deviation of the fgt and bmi. The rule base has around 40 rules; one of them is shown below.

```
3.        if (fgt==’low’ && bmi==’low’)
              then tid=’low’;
          else if (fgt==’low’ && bmi==’high’)
              then tid=’high’;
```

II. Type-1 FL based insulin dosage recommendation: Fuzzification of the input is performed when a valid input is received, as shown in (4).

a) The type-1 fuzzy based inference system loads the rule base and matching degree is calculated for each rule. As we have three consequences of final tid, i. e. ‘low’ dosage, ‘high’ dosage, ‘med’ dosage, three sets of the fired rules will be formed.

b) The fuzzy MAX disjunction is calculated for each consequence, i.e. the rule with maximum matching degree and stored in y (i), where i is the no. of consequences (3 in this case).

c) Defuzzification is applied to these three outputs using the centroid method and the result shows the insulin dosage in unit/kg/day for a single day. As shown in (7), *w (i)* is the weight of aggregated rules.

III. Semantic description of the output: The probability so calculated is arranged as shown in table 5.2. The final output shows bmi and fgt, according to their semantic description.

4.6 Interval Type-2 FL based insulin dosage prediction

The Interval Type-2 FL can also be used for insulin dosage prediction like Type-1 FL. The knowledge base is based on the domain of insulin dosage.

I. Knowledge base: The knowledge base consists of the Type-2 FL based rule base and a fuzzified insulin dataset. The rule base has ‘tid’ as the class variable, which can attain any value from 0 to 1. The rule base has around 40 rules; one of them is shown below.

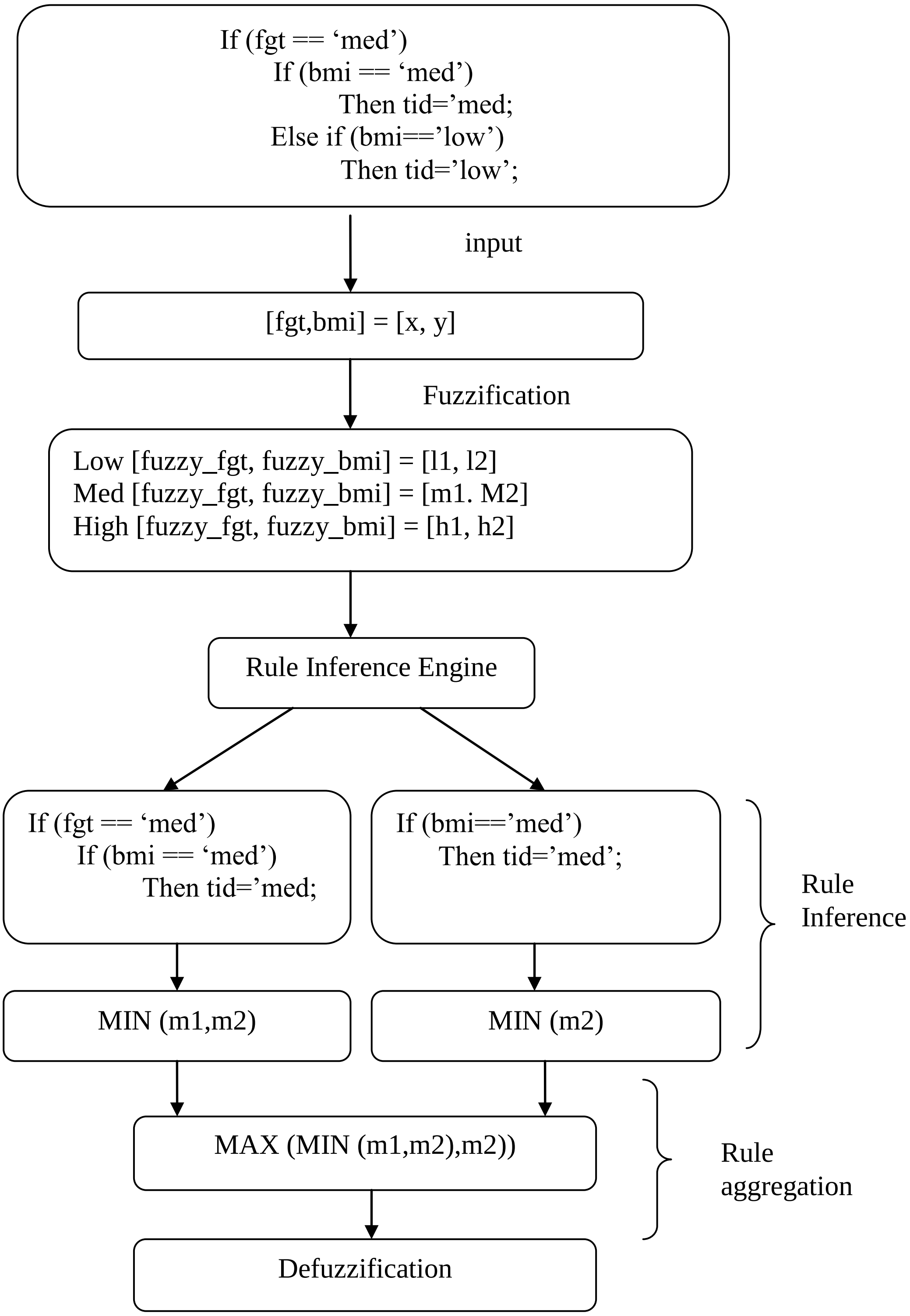

Figure 4.4 Insulin Dosage prediction using Type-1 FL

II. Interval Type-2 FL Insulin Dosage prediction:

a) Fuzzification: FoU is calculated for the both the attributes, i.e. bmi and fgt as shown in (8), (9) and (10). C is chosen in order to keep probability in acceptable limits. So, it is preferred to keep C=0.2 or C=0.5. Then trapezoidal membership function is used to compute the upper, lower and T1FL based secondary fuzzy values.

b) Rule Inference: Rule inference is similar to that of type-1 fuzzy logic. Rule inference uses the MIN fuzzy based conjunction operation. The resulting fuzzy rule inference value is stored in different sets according to the class variable of the rule as shown in figure 4.4.

c) Type-reduction: Left and right switch points are calculated using (14) and (15) respectively, for the fired rules of each consequence. These switch points are then used to calculate rule aggregation values for each consequence.

d) Defuzzification: Using the centroid method, the crisp output, i.e. the probability can be obtained.

III. Semantic decision making: As shown in table 5.2, the insulin dosage is interpreted as 'low', 'med' and 'high'. The output shows the fuzzy sets of the bmi and fgt just like in the case of Type-1 FL based output.

4.7 Mathematical Model

I. Computation of fuzzy numbers: For fuzzification, it is important to define the fuzzy numbers. In the thesis, we are using three fuzzy numbers for each attribute, i.e. att_low, att_med, att_high as shown below:

att_low: *m=min(att), n=min(att), o=mean(att)-stddev(att), p = mean(att)*
att_med: *m=mean(att)–stddev(att), n=mean(att), o= mean(att), p=mean(att)+stddev(att)*
att_high: *m=mean(att), n=mean(att)+stddev(att), o =max(att), p=max (att)*

Each fuzzy number defines some area of the membership function as low, med and high for a particular value depending on its minimum, maximum, mean and standard deviation.

II. Uncertainty constant C: The constant c is calculated to compute the upper and lower membership function of an attribute while decision making. C converts the irregular FoU to regular FoU as it helps in computing the upper and lower membership functions. Rather than randomly choosing C as any number between 0 and 1, for better accuracy it is preferred to be calculated as shown in (13).

$$C = (\sum_{i=1}^{n} \mu_{\tilde{A}})/ n \tag{13}$$

Where, n = No. of the rules fired by type-1 fuzzy logic

$\mu_{\tilde{A}}$ = Membership function of an attribute 'A'

In this way, C captures the average value so that not only it captures value between 0 and 1, but is also confined to the range of the membership function values being calculated for that particular input.

III. Calculation of right and left switch points in Type-2 FL: The calculation of the switch points is in fact the most important one. As shown below in (14) and (15), while using a trapezoidal function, the left and right switch points depend on the N. With the increase in the number of fired rules, i.e. N, the no. of switch points increase over the membership function. In (14) and (15), the $\underline{T}$ is the upper membership function value and $\overline{T}$ is the lower membership function.

$$\underline{T} = N/2.4 \tag{14}$$

$$\overline{T} = N/1.7 \tag{15}$$

$$\overline{M}_j = \sum_{i=1}^{T_i} x_i * \overline{I}_{i,j} \ + \sum_{i=\underline{T}_i+1}^{n} x_i * \underline{I}_{i,j} \,/ \sum_{i=1}^{T_i} \overline{I}_{i,j} \ + \sum_{i=\underline{T}_i+1}^{n} \underline{I}_{i,j} \tag{16}$$

$$\underline{M}_j = \sum_{i=1}^{\overline{T}_i} x_i * \underline{I}_{i,j} \ + \sum_{i=\underline{T}_i+1}^{n} x_i * \overline{I}_{i,j} \,/ \sum_{i=1}^{T_i} \underline{I}_{i,j} \ + \sum_{i=\underline{T}_i+1}^{n} \overline{I}_{i,j} \tag{17}$$

$\overline{M}_j$ is the left switch value,

$\underline{M}_j$ is the left switch value, where j is the type of the consequence

x_i is the secondary membership function value or type-1 fuzzification base value,

$\overline{I}_{i,j}$, I is the rule inference value and $\overline{I}$ is the value of the rule inference in the upper FOU .

Here left and right switch points in the FoU split the rules into two sets, first set of which the upper matching degree is considered and the second set for which the lower rule matching degrees are considered. It is similar to rule aggregation of the type-1 fuzzy logic.

5. SIMULATION AND RESULTS

5.1 Evaluation Parameters

I. Probability: The prediction of the disease diagnosis can be very low, low, med, high and very high, as shown in table 5.1.

Table 5.1 Semantic description of the probability of the disease diagnosis

Probability	Semantic description
0 – 0.20	very low
0.21 – 0.40	low
0.41 – 0.60	med
0.61 – 0.80	high
0.80 - 1	very high

The attributes are analyzed on the basis of their fuzzy values for the formation of the rule base and for computing the matching degree. Similarly for insulin dosage, different semantic descriptions have been set for the attributes as shown in table 5.2.

Table 5.2 Semantic description of different levels of insulin dosage

Probability of Insulin Dosage	Semantic description
0 - 0.3	low
0.3 - 0.6	med
0.6 - 1	high

II. Rule base: As in the proposed work, two methodologies i.e. Type-1 FL and Interval Type-2 FL are used and there are differences in their respective rule bases. In latter, the rule base includes the upper and lower fuzzy values of the attributes, making it much larger than the rule base of Type-1 FL.

III. Comparison of the final output: The comparison is drawn on the basis of the difference of the outputs of Type-1 FL and Interval Type-2 FL for both the diseases to understand three things:

i. The difference of the two methods by taking into account the predictions made by them.

ii. Analyzing whether the difference is uniform or non-uniform with the changing stages of the symptoms.

iii. The applicability of I-T2FL in medication where domain knowledge is much smaller in size than the diagnosis dataset.

5.2 Results

1. For input I = [165, 80, 12, 11, 23, 45, 0.2], the output of the diabetes diagnosis prediction by Type-1 Fuzzy Logic is shown below.

> fgt is high, sit is low, bmi is low, age is low and df is high. Type 1 FL Probability=0.9159. So, chances are very high.

2. Table 5.3 depicts the fuzzy numbers, i.e., age_low, age_med, age_high for age attribute of heart related complications.

Table 5.3 [m,n,o,p] set for trapezoidal membership for age attribute for heart related complications prediction

Fuzzy numbers	**[m,n,o,p]**
age_low	[34, 34, 49.59, 55.725]
age_med	[49.59, 55.725, 55.725, 61.85]
age_high	[55.725, 61.85, 74, 74]

3. For input [age, gender, cp, trestbps, thalach, exang, oldpeak, slope] = [56, 1, 2, 100, 80, 0.33, 2, 1], the type-1 fuzzy logic based prediction of heart related complications is:

> age is medium, cp is low, trestbps is low, thalach is low, exang is medium, oldpeak is high, slope is low. Type 1 Probability=0.7903743916549264. So, chances are high.

4. For input, pgc =140, si=400, bmi=23, mbp=25, age=56, df= 0.55. Type-2 fuzzy logic based diabetes diagnosis probability is:

> pgc is high, si is high, bmi is low, mbp is medium, age is high. Probability=0.5 So, chances are medium. Type 2 Probability=0.310. So, chance is low.

5. For input, age = 56, gender = 1 (male), cp = 3, trestbps = 100, thalach = 100, exang = 0, oldpeak = 0, slope = 2, the heart related complications prediction based on T2FL and T1FL is:

> age is medium, cp is low, trestbps is low, thalach is low, exang is low, oldpeak is medium. Type 1 probability=0.9999999999999999. So, the chances are very high. Type 2 Probability=0.5681006902004474. So, chances are medium.

6. Table 5.4 shows the fuzzy numbers calculated for the fgt while predicting diabetes diagnosis based on type-1 fuzzy logic.

Table 5.4 Fuzzy numbers of fgt for insulin dosage prediction

Fuzzy numbers	**[m,n,o,p]**
Fgt_low	[0, 0, 98.30119, 120.89453]
Fgt_med	[98.30119, 120.89453, 120.89453, 143.48787]
Fgt_high	[120.89453 143.48787 199 199]

7. For the input [fgt, bmi] = [106, 30.5], the output for insulin dosage according to type-1 fuzzy logic is

> fgt is low, bmi is med. Type-1 fuzzy dosage = 0.3201, i.e., ‘med’ level of dosage.

8. Table 5.5 shows the three dimensional fuzzy logic based fuzzy numbers stored in the knowledge base.

Table 5.5 Upper, lower and secondary fuzzy numbers for diabetes for input [pgc, si, bmi, dpf, age] = [136, 74, 50, 204, 37.4, 28]

Attribute	Fuzzy numbers	Type-1 fuzzy value	Lower membership function	Upper membership function
fgt	**fgt_low**	0.0	0.0	0.25
	fgt_med	0.346	0.096	0.596
	fgt_high	0.653	0.403	0.903
sit	**sit_low**	0.0	0.0	0.25
	sit_med	0.0	0.0	0.25
	sit_high	1.0	0.5	1.0
bmi	**bmi_low**	0.0	0.0	0.25
	bmi_med	0.081	0.0	0.268
	bmi_high	0.981	0.5	1.0
df	**df_low**	0.311	0.061	0.561
	df_med	0.688	0.438	0.938
	df_high	0.0	0.0	0.25
age	**age_low**	0.642	0.392	0.892
	age_med	0.357	0.107	0.607
	age_high	0.0	0.0	0.25

9. For the input [fgt, bmi] = [106, 30.5], the output for insulin dosage according to type-2 fuzzy logic is

fgt is low, bmi is med. So, type-2 fuzzy dosage = 0.2283, i.e., 'low' level of dosage

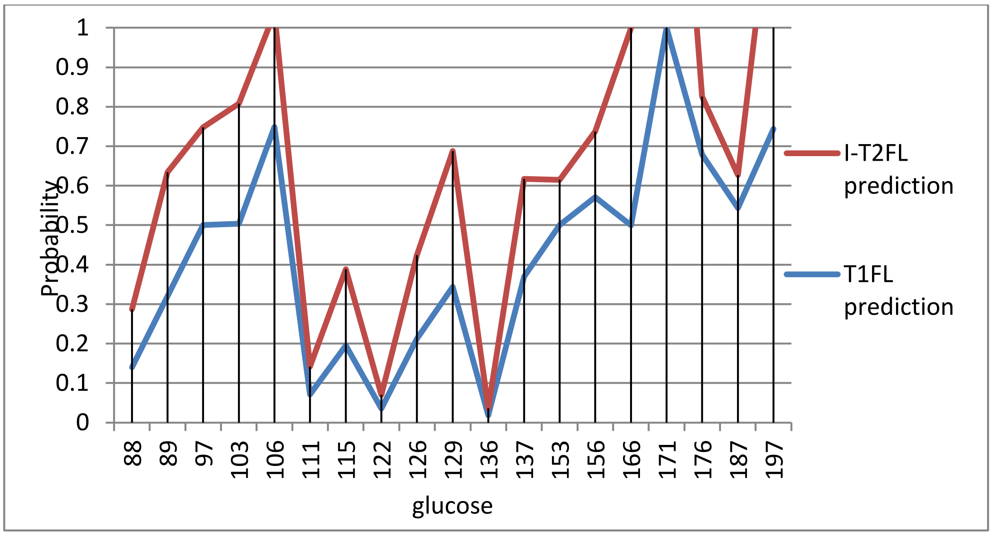

Figure 5.1 Graph depicting differences of the probabilities calculated by T1FL and I-T2FL

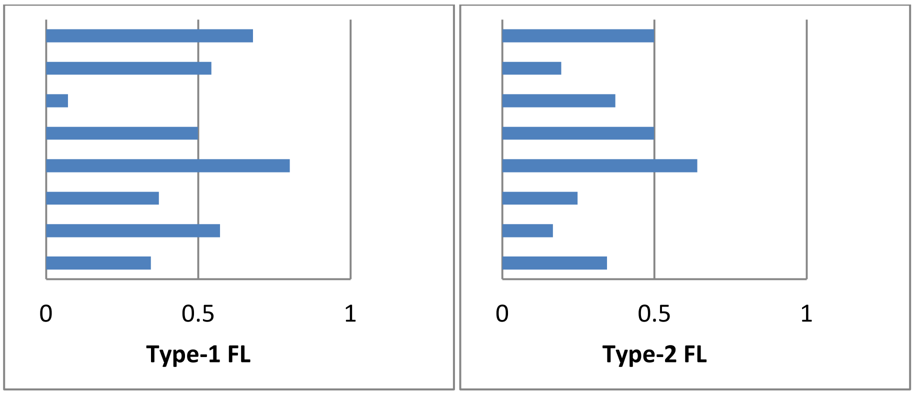

Figure 5.2 Predictions by Type-1 FL and Type-2 FL for positive diabetes diagnosis

I. The comparison of Type-1 FL and Interval Type-2 FL for type-2 diabetes diagnosis prediction: Figure 5.1 shows the difference of the outputs calculated for 20 inputs for diabetes.

a) There is no constant, but variable gap between the Type-1 FL and Interval Type-2 FL based diagnosis prediction as the values of symptoms change.

b) The probability of diagnosis calculated by Interval Type-2 FL is always greater than that of the Type-1 FL.

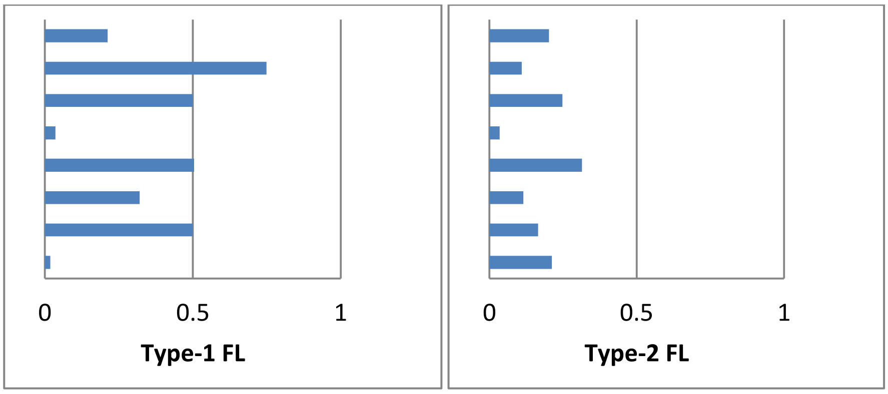

Figure 5.3 Predictions by Type-1 FL and Type-2 FL for negative diabetes diagnosis

From figure 5.2, it is noticeable that the predictions made by both approaches are nearly same in case of positive diagnosis of diabetes. However, in the case of negative diagnosis, the picture is quite different as shown in fig 5.3. Type-2 Fuzzy Logic attempts to achieve better accuracy (nearly 90%), where Type-1 Fuzzy Logic lags and achieve accuracy up to just around 75%.

II. The comparison on the basis of heart related complication prediction

The figure 5.4 shows the gap between the predictions made by the two methodologies.

a) It is noticeable that unlike in the case of diabetes, Type-1 FL always predicts greater probability than Interval Type-2 FL.

b) The gap between the two predictions is more uniform than that in the case of diabetes diagnosis prediction.

c) In case of heart diseases, Type-1 FL outperforms Interval Type-2 FL proving itself to be a better approach for diagnosis of heart related complications.

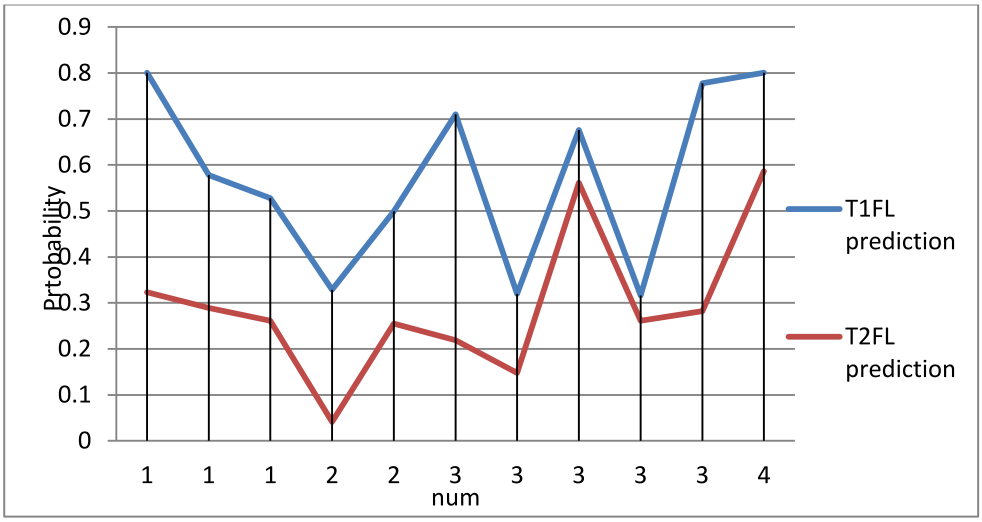

Figure 5.4 Graph depicting the probabilities calculated by T1FL and I-T2FL with respect to num

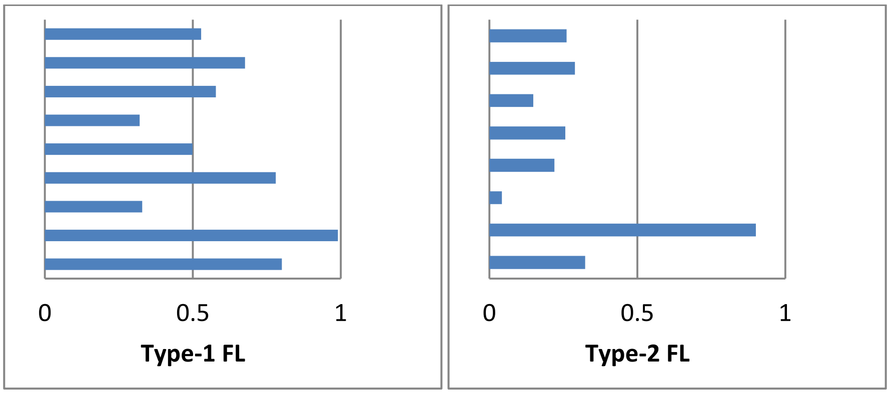

Figure 5.5 Probabilities calculated by Type-1 FL and Interval Type-2 FL for positive diagnosis of heart diseases

III. Comparison of T1FL and I-T2FL in insulin dosage prediction

Figure 5.6 shows the expanding gap between the predictions of the two approaches. In the case of insulin, both the methodologies achieve nearly the same accuracy rate of 80%, and it is here where fuzzy logic can be used with other soft computing techniques to achieve better accuracy.

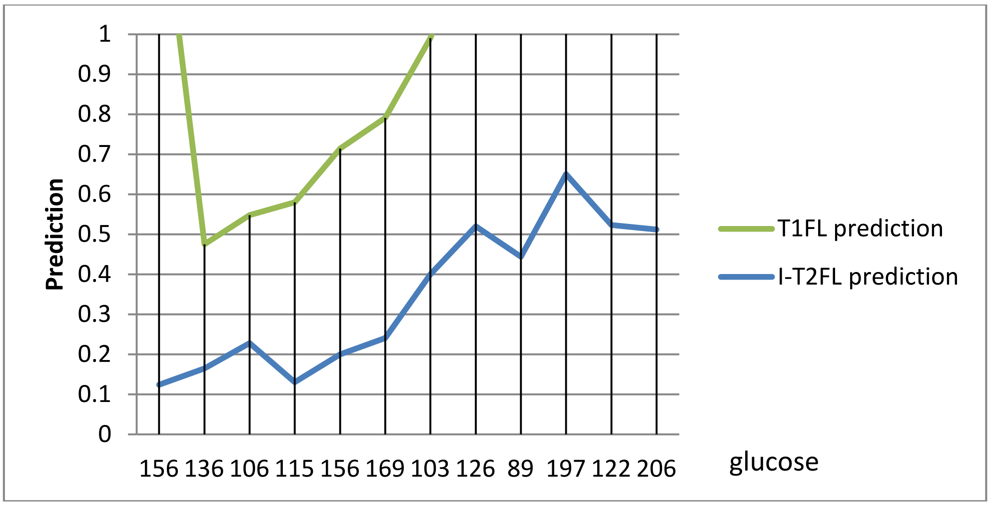

Figure 5.6 Graphical comparison of insulin dosage prediction

6. CONCLUSION AND FUTURE SCOPE

6.1 Conclusion

The performance of Type-1 FL and interval Type-2 FL is different with respect to different diseases. It challenges the norm that the latter is always better than the former. The complexity of diabetes and heart related disease diagnosis is almost equal, yet the difference in the outputs of Type-1 FL and Type-2 FL triggers the need to further study these approaches in medical diagnosis.

For medication recommendation, the Interval Type-2 FL is useful if the rule base is significantly larger and contains distinct rules with high confidence. And if the dataset is small like the diabetes medication data, Type-1 FL is a preferred approach.

6.2 Future Scope

The comparison of Type-1 FL and Interval Type-2 FL is presented only for diabetes and heart related complications which can also be performed on other diseases like HIV and TB. Due to non- availability of a universal dataset of these two diseases, the research in these diseases using fuzzy is quite limited. Besides this, various other metrics like complexity and time taken for diagnosis can be used for the comparison of the two approaches.

It is worth finding if Type-1 FL and Interval Type-2 FL can be used together in collaboration for better results and optimized use of resources. In medication, as both the approaches are unable to achieve a higher rate of accuracy, better approaches can be formulated by fusing these approaches with other soft computing methodologies.

REFERENCES

[1] C. Yau and A. Sattar, "Developing Expert System with Soft Systems Concept," *In proceedings of International Conference on Expert Systems for Development*," pp. 79-84, 1994.

[2] N. Conigliaro, A. Stefano, and O. Mirabella, "An Expert System for Medical Diagnosis," *In proceedings of Symposium on the Engineering of Computer-Based Medical Systems*, pp. 75-81, 1988.

[3] K. Levi, "Expert Systems Should be More Accurate than Human Experts: Evaluation Procedures from Human Judgment and Decision making," *IEEE Transaction on Systems, Man, and Cybernetics*, vol. 19, no. 3, pp. 647-657, 1989.

[4] I. M. Ahmed, A. M. Mahmoud, M. Aref, and A. M. Salem, "A study on Expert Systems for Diabetic Diagnosis and Treatment," *Recent Advances in Information Science*, pp. 363-367, 2012.

[5] S. Rahaman, "Diabetes Diagnosis Decision Support System based on Symptoms, Signs and Risk Factor using Special Computational Algorithm by rule Base," *In proceedings of 15th International Conference on Computer and Information Technology (ICCIT)* , pp. 65-71, 2012

[6] J. J. Kim and G. A. Bekey, "Adaptive Abstraction in Expert Systems For Medical Diagnosis," *In Proceedings of Fifth Annual IEEE Symposium on Computer-Based Medical Systems*, pp. 345-352, 1992.

[7] K. V. Lakshmi and M. Padmavathamma, "Modeling an Expert System for Diagnosis of Gestational Diabetes Mellitus Based On Risk Factors," *IOSR Journal of Computer Engineering*, vol. 8, no. 3, pp 29-32, 2013.

[8] H. Temurtas, N. Yumusak, and F. Temurtas, "A comparative study on diabetes disease diagnosis using neural networks," *Expert Systems with Applications*, vol. 36, no. 4, pp. 8610-8615, 2009.

[9] R. Dey, V. Bajpai, G. Gandhi, and B. Dey, "Application of Artificial Neural Network (ANN) technique for Diagnosing Diabetes Mellitus," *In proceedings of IEEE Third international Conference on Industrial and Information Systems*, pp. 1-4, 2008.

[10] A. M. Aibinu, M. J. E. Salami, and A. A. Shafie, "Application of Modeling Techniques to Diabetes Diagnosis," *In proceedings of IEEE EMBS Conference on Biomedical Engineering & Sciences (IECBES),* pp. 194-198, 2010.

[11] C. Lee and M. Wang, "A Fuzzy System for Diabetes Decision Support Application", *IEEE Int. Conference on Systems, Man, and Cybernetics,* vol. 41, no. 1, pp. 139-153, 2010.

[12] S. Calegari and E. Sanchez, "A Fuzzy Ontology-Approach to improve Semantic Information Retrieval," *URSW,* vol. 327, pp. 1-6, 2007.

[13] D. N. M. Abadi, M. H. Khooban, and Mehdi Siahi, "A Novel Automated Fuzzy Model for Diabetes Mellitus," *IEEE 2nd International Conference on Control, Instrumentation and Automation (ICCIA),* pp. 350-354, 2011.

[14] I. Gadaras and L. Mikhailov, "An interpretable fuzzy rule-based classification methodology for medical diagnosis," *Artificial Intelligence in Medicine,* vol. 47, no. 1, pp. 25-41, 2009.

[15] R. R. Yager and F. E. Petry, "A Multicriteria Approach to Data Summarization Using Concept Ontologies," *IEEE Transactions on Fuzzy Systems*, vol. 14, no. 6, pp. 767-780, 2006.

[16] J. Palma, J. M. Juarez, M. Campos, and R. Marin, "Fuzzy theory approach for temporal model-based diagnosis: An application to medical domains," *Artificial Intelligence in Medicine,* vol. 38, no. 2, pp. 197—218, 2006.

[17] R. Seising, "From vagueness in medical thought to the foundations of fuzzy reasoning in medical diagnosis," *Artificial Intelligence in Medicine*, vol. 38, no. 3, pp. 237-256, 2006.

[18] E. Straszecka, "Combining uncertainty and imprecision in models of medical diagnosis," *Information Sciences,* vol. 176, no. 20, pp. 3026-3059, 2006.

[19] M. Kalpana and A. V. S. Kumar, "Fuzzy Expert System for Diagnosis of Diabetes Using Fuzzy Determination Mechanism," *International Journal of Computer Science and Emerging Technology*, vol. 2, no. 6, pp. 39-45, 2011.

[20] S. Das and S. Kar, "Group decision making in medical system: An intuitionistic fuzzy softest," *Applied Soft Computing*, vol. 24, pp. 196-211, 2014.

[21] M. Mahfouf, M. F. Abbod, and D. A. Linkens, "A survey of fuzzy logic monitoring and control utilisation in medicine," *Artificial Intelligence in Medicine,* vol. 21, no. 1-3, pp. 27-41, 2001.

[22] P. Innocent and R. John, “Computer aided fuzzy medical diagnosis,” *Information Sciences*, vol. 162, pp. 81-104, 2004.

[23] H. Kahramanli and N. Allahverdi, “Design of a hybrid system for the diabetes and heart diseases,” *Expert Systems with Applications*, vol. 35, no. 1-3, pp. 82–89, 2008.

[24] N. K. Gupta, A. Gupta, and P. K. Tyagi, “Early Detection of Diabetes Patients using Soft Computing,” *In proceedings of International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*, pp. 174-179, 2014.

[25] S. M. Anouncia, L. J. C. Madonna, P. Jeevitha, and R. T. Nandhini, “Design of a Diabetic Diagnosis System Using Rough Sets,” *Cybernetics and Information Technologies*, vol. 13, no. 3, pp. 124-139, 2013.

[26] Chang-Shing Lee, Mei-Hui Wang, and H. Hagras, “A Type-2 Fuzzy Ontology and Its Application to Personal Diabetic-Diet Recommendation,” *IEEE Transactions on Fuzzy Systems*, vol. 18, pp. 374-395, 2010.

[27] J. M. Mendel, H. Hagras, and R. I. John, “Guest Editorial for the Special Issue on Type-2 Fuzzy Sets and Systems,” *IEEE Transactions on fuzzy systems*, vol. 21, no. 3, pp. 397-398, 2013.

[28] C. Yeh, W. R. Jeng, and S. Lee, “An Enhanced Type-Reduction Algorithm for Type-2 Fuzzy Sets,” *IEEE Transactions on fuzzy systems*, vol. 19, no. 2, pp. 227-240, 2011.

[29] J. H. Aladi, C. Wagner, and J. M. Garibaldi, “Type-1 or INTERVAL Type-2 Fuzzy Logic Systems - On the Relationship of the Amount of Uncertainty and FOU Size,” *In Proceedings of IEEE Conference on Fuzzy Systems*, pp. 2360-2367, 2014.

[30] A. Roychowdhury, D. K. Pratihar, N. Bose, K. P. Sankaranarayanan, and N. Sudhahar, “Diagnosis of the diseases—using a GA-fuzzy approach,” *Information Sciences*, vol. 162, no. 2, pp. 105-120, 2004.

[31] S. H. Ling, H. Nguyen, and K. Y. Chan, “Genetic Algorithm based Fuzzy Multiple Regression for the Nocturnal Hypoglycaemia Detection,” *In proceedings of IEEE World Congress on Computational Intelligence*, pp. 18-23, 2010.

[32] K. Y. Chan, S. H. Ling, T. S. Dillon, and H. Nguyen, "Classification of Hypoglycemic Episodes for Type 1 Diabetes Mellitus based on Neural Networks," *In proceedings of IEEE Congress on Evolutionary Computation*, pp. 1-5, 2010.

[33] Yi-Zeng Hsieh, Mu-Chun Su, and Pa-Chun Wang, "A PSO-based rule extractor for medical diagnosis," *Journal of Biomedical Informatics,* vol. 49, pp. 53–60, 2014.

[34] H. H. Inbarania, A. T. Azarb, and G. Jothi, "Supervised hybrid feature selection based on PSO and rough sets for medical diagnosis," *Computer Methods and Programs in Biomedicine*, vol. 113, no. 1, pp. 175-185, 2014.

[35] M. Kumar, A. Sharma, and S. Agarwal, "Clinical Decision Support System for Diabetes Disease Diagnosis Using Optimized Neural Network," *In proceedings of Students Conference on Engineering and Systems (SCES)*, pp. 1-6, 2014.

[36] M. F. Ganji, M. S. Abadeh, "Using fuzzy Ant Colony Optimization for Diagnosis of Diabetes Disease," *In Proceedings of Iranian Conference of Electrical Engineering*, pp. 501-505, 2010.

[37] S. Kumari and A. Singh, "A Data Mining Approach for the Diagnosis of Diabetes Mellitus," *In Proceedings of 7th international Conference on Intelligent Systems and Control*, pp. 373-375, 2013.

[38] B. M. Patil, R. C. Joshi, and Durga Toshniwal, "Association rule for classification of type-2 diabetic patients," *In proceedings of 2010 Second International Conference on Machine Learning and Computing*, pp. 330-334, 2010.

[39] M. Kwiatkowska and K. Kielanb, "Fuzzy logic and semiotic methods in modeling of medical concepts," *Fuzzy Sets and Systems,* vol. 214, pp. 35-50, 2013.

[40] S. C. Suh and G. P. Vudumula, "The Role of Conceptual Hierarchies in the Diagnosis and Prevention of Diabetes," *In proceedings of 7th International Conference on Networked Computing and Advanced Information Management (NCM)*, pp. 267-275, 2011.

[41] N. Esfandiari, M. Babavalian, A. E. Moghadam, and V. K. Tabar, "Knowledge discovery in medicine: Current issue and future trend," *Expert System with Applications*, vol. 41, no. 9, pp. 4434-4463, 2014.

[42] D. Campos-Delgado, M. Hernández-Ordoñez, R. Femat, and A. Gordillo-Moscoso, "Fuzzy-Based Controller for Glucose regulation in Type-1 Diabetic Patients by Subcutaneous Route," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 11, pp. 2201-2209, 2006.

[43] P. Grant, "A new approach to diabetic control: Fuzzy logic and insulin pump technology," *Medical Engineering & Physics*, vol. 29, no. 7, pp. 824-827, 2007.

[44] D. Nazari, M. Abadi, M. H. Khooban, and Mehdi Siahi, "A Novel Automated Fuzzy Model for Diabetes Mellitus," *In proceedings of IEEE 2nd International Conference on Control, Instrumentation and Automation (ICCIA)*, pp. 350-354, 2011.

[45] A. Adeli, M. Neshat, "A Fuzzy Expert System for Heart Disease Diagnosis," *In Proceedings of International MultiConference of Engineers and Computer Scientists*, vol. 1, pp. 1-6, 2010.

LIST OF PUBLICATIONS

[1] N. Lalka, S. Jain, " Fuzzy Based Expert System for Diabetes Diagnosis and Insulin Dosage Control", *In proceedings of International Conference on Computing Communication and Automation (ICCCA-2015),* 2015. (ACCEPTED and PRESENTED)

[2] Comparative Study of Type-1 Fuzzy Logic and Type-2 Fuzzy Logic (to be confirmed).

VIDEO PRESENTATION

Video link:
https://www.youtube.com/watch?v=qLYV23KatYg&feature=youtu.be

Slide 1: Student name, name of the topic and Institution name
Slide 2: Topics to be discussed
Slide 3: Medical diagnosis
Slide 4: Fuzzy Logic
Slide 5: Type-2 Fuzzy Logic
Slide 6: Difference of the working of T1FL and T2FL
Slide 7: Benefits to the society

REFLECTIVE DIARY

An ideal learning process always incorporates ups and downs. In fact lessons come our way unexpectedly, and that is why it is said ‘To be prepared is half the victory’. First of all, I think Dr. Sushma Jain as my guide did wonders to my year long learning process. She had that helping hand, which I needed when I often met the ground. And as it was the first time I was to do something called ‘research’, I did not know even a bit about researching like an expert. There were many incidents when I had to stop and think.

1. August, 2014 – September, 2014

Challenge: Developing an art of grasping the relevant material from a research paper.

I had made up my find about pursuing my thesis in Artificial. There was a choice of either working with AI in robotics (I didn’t quite know about what exactly it meant) or medical diagnosis. AI in Medical Diagnosis is something that is hotly discussed over blogs and other websites. In this way, I chose this sub topic as there was more availability of data sets and sources to understand medical diagnosis.

I studied Soft Computing in my course earlier, so it wasn’t that difficult to grab the sources. What was again a challenge was ‘Where to start from?’ Reading papers is often an answer, but being a newbie that took me like a month or two to figure out what paper or research work was useful to me. Later choosing a sub topic became the next challenge.

2. October, 2014 – December, 2014

Challenge: Understanding fuzzy logic

In October, I was clear about doing something with the Fuzzy Logic. And as it was told to us that mid-November may be the time when we have to present our respective thesis work in front of the teachers and guides. The challenge that I met here was forming a schedule about when I wanted to do an implementation. But with the guidance of my guide, I was able to concentrate on finding a proper sub topic and reading papers about it. Yes, I was late enough to figure out the sub topic. (Some of my friends were clear from the very start).

In November, I started implementing fuzzy logic on diabetes data set. So, it was ‘Fuzzy Logic for diabetes diagnoses’, something incomplete, I would say. But this is how things start. I was ready with almost everything till my presentation (which was postponed to December 16), the only problem remained ‘how to present like a researcher?’ Various dry things were there in the presentation. I learnt a lot from the brainstorming during the presentation, it gave me a confidence about answering things about which I was the only student in the class to know (No one liked fuzzy logic during the course either). However, I would not say the presentation went the way I wanted. It was another lesson, I was completely a newbie.

3. January, 2015 – February, 2015

Challenge: Understanding insulin dosage patterns for using them in type-1 fuzzy logic

I had to take my work bit further by formulating an expert system for medication. It was also the time when we were supposed to submit our respective work for conferences. I could hardly gather real time data, so I relied on the data set from the internet. And in this way, I was ready with ‘Fuzzy Logic on diabetes and insulin dosage recommendation’. I submitted my paper and later, after acceptance went to present it. It was a new experience. What became the source of inspiration were my knowledge and my complete understanding of the work that I did. It was one of the most useful learning experiences. Experts were sharing every bit of suggestion that they could give and it is where I thought of expanding my thesis work, implementing something which is less researched about.

4. March, 2015 – April, 2015

Challenge: Implementing Type-2 fuzzy logic

I was on my way to implement Type-2 Fuzzy logic, an advanced form of fuzzy logic. It is the most prominent thing about what I did on my thesis. Being a mathematically tougher approach, it took me longer than desired to learn it. Here, time management was a challenge, as on one way we were to compile things and start writing a thesis. The Karnik-Mendel algorithm took me long to implement it on medical data.

5. May, 2015 – June, 2015

June, a month that was full of challenges. Often one would consider it to be the most relaxing time, but as per my experience it was an even more tiring process than the implementation. Thesis writing as I learnt, requires every single piece of the work to be penned down. Every single bit of the way things are done during the thesis has to be written. What if I would forget mentioning something I spent days working on? So, it is the thesis writing, which I think is one of the most important part of the project work.

In this way, I would like to sum up my whole experience which triggered a thought process and anxiety to learn more about computer science.