Data Modeling with Type I and Type II Fuzzy Sets

Zina Bibo Bilasini

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Computer Engineering

Eastern Mediterranean University
September 2016
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Mustafa Tümer
Acting Director

I certify that this thesis satisfies the requirements as thesis for the degree of Master of Science in Computer Engineering.

Prof. Dr. Işık Aybay
Chair, Department of Computer Engineering

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Master of Science in Computer Engineering.

Asst. Prof. Dr. Adnan Acan
Supervisor

Examining Committee

1. Asst. Prof. Dr. Adnan Acan
2. Asst. Prof. Dr. Mehmet Bodur
3. Asst. Prof. Dr. Ahmet Ünveren

ABSTRACT

The fuzzy classifier is an algorithm that assigns a class label to an object, based on the object description. It is also said that the classifier predicts the class label. The object description comes in the form of a vector containing values of the features (attributes) deemed to be relevant for the classification task. Typically, the classifier learns to predict class labels using a training algorithm and a training data set. When a training data set is not available, a classifier can be designed from prior knowledge and expertise. Once trained, the classifier is ready for operation on unseen objects.In this thesis, type-1fuzzy classifier, and the type-2 fuzzy classifier are used for the machine learning datasets classification. The Wisconsin breast cancer dataset, Iris Dataset, and Tic-Tac-Toe datasets are classified. Type-2 fuzzy classifiers are able to perform better than type-1 fuzzy classifiers which have additional design parameters. Therefore, type-2 fuzzy classifiers are more attractive than the type-1 classifiers. The essential benefits the type-2 fuzzy logic classifiers are their ability to handle more vagueness.

Keywords: Classifier, Type-1fuzzy classifier, Type-2 fuzzy classifier, Machine learning dataset and Uncertainty.

ÖZ

Bulanık sınıflayıcı verilen bir nesneler kümesindeki her eleman için nesne tanımına göre bir sınıf etiketi atayan bir algoritmadır. Sınıflayıcı için aynı zamanda sınıf etiketini belirleyici denilir. Nesne tanımı bir dizi içerisinde sınıflandırma işlemiyle ilgili nesne özelliklerinin verilmesiyle yapılır. Tipik olarak, sınıflayıcı bir öğrenme kümesi üzerinden bir öğrenme algoritması kullanarak sınıf etiketlerini öğrenir. Öğrenme kümesinin varolmadığı durumlarda, sınıflayıcılar tecrübeye ve uzman bilgisine göre tasarlanırlar. Öğrenme aşamasının ardından, sınıflayıcı daha önce karşılaşmadığı nesneleri sınıflamaya hazır olur. Bu tezde, tip-1 ve tip-2 bulanık sınıflayıcılar makine öğrenme veri-kümeleri üzerinde sınıflayıcı olarak kullnılmışlardır. Wisconsin göğüs kanseri, Iris ve tic-tac-toe oyunu veri kümelerindeki nesneler deneysel çalışmalarda sınıflandırılmışlardır. Tip-2 bulanık sınflayıcıların başarımının tip-1 bulanık sınıflayıcılara göre daha iyi olduğu gözlemlenmiştir. Tip-2 bulanık sınıflayıcıların daha iyi başarım göstermesinin nedeni, daha fazla tasarım parametresine sahip olmaları ve böylece belirsizliği daha iyi modellemeleri olarak belirtilebilir.

Anahtar Kelimeler: Sınıflayıcı, Tip-1 Bulanık, Tip-2 Bulanık, Makine Öğrenme Veri-Kümeleri and Belirsizlik.

DEDICATION

To my beloved family

&

My best friends

ACKNOWLEDGMENT

In the name of Allah, the Most Gracious and the Most Merciful. All praises to Allah for the strengths and his blessing in completing this thesis.

It gives me great pleasure in acknowledging the support and help of my supervisor, Asst.Prof.Dr. Adnan Acan for his invaluable guidance, supervision, encouragement and constant support during the course of this research.

I express my warm thanks to my best friends Lutfia,Basma,Vaman Saeed and Amar Mustafa who always help me.

I wish to express my love and gratitude to my beloved families for their understanding and endless love through the duration of my studies.

TABLE OF CONTENTS

ABSTRACT iii

ÖZ iv

DEDICATION v

ACKNOWLEDGMENT vi

LIST OF TABLES x

LIST OF FIGURES xii

LIST OF ABBREVIATIONS xiii

1 INTRODUCTION 1

1.1 Fuzzy Logic – an Overview 1

1.2 Fuzzy Model 1

1.3 Fuzzy Classification 2

1.4 Fuzzy C-Means Clustering 2

1.5 Problem Statement 3

1.6 Significance of Research 3

1.7 Thesis Organization 4

2 LITERATURE REVIEW OF THE RELEVANT RESEARCH 5

3 FUNDAMENTALS OF TYPE-1 AND TYPE-2 FUZZY LOGIC 8

3.1 Fuzzy Set Theory 8

3.2 Properties of Fuzzy Sets 9

3.3 Membership Function of a Fuzzy Set 10

3.3.1 Membership Function Terminology 11

3.4 Type-2 Fuzzy Sets and Systems 11

3.4.1 Fuzzy Sets and Uncertainty 12

3.4.2 Type-2 Fuzzy Sets 13

3.4.3 Operations of Type-2 Fuzzy Sets 14

3.5 Fuzzy Inference System 16

3.5.1 The Architecture of Fuzzy Inference Systems 17

3.5.2 Fuzzy Rule Based System 19

3.5.3 Mamdani Fuzzy Rule Based System 19

4 IMPLEMENTATION OF THE PROPOSED SYSTEM 21

4.1 Introduction 21

4.2 Methodology 21

4.3 Principles of Fuzzy Classifier 23

4.3.1 Type-1 Classifier General Process Steps 23

4.3.2 Type-2 Classifier Procedural Steps 25

4.4 Fuzzy Type-1 Classifier Algorithm 27

4.5 Fuzzy Type-2 Classifier Algorithm 29

4.6 Evaluate The Fuzzy Classification 31

5 RESULTS AND DISCUSSION 32

5.1 Experiments and Results 32

5.2 Wisconsin Breast Cancer Data Set 33

5.3 Iris Flower Data Set Description 34

5.4 Tic-Tac-Toe Endgame Data Set 35

5.5 Fuzzy Type-1 Classifier Result on Breast Cancer Dataset 36

5.6 Fuzzy Type-1 Classifier Result on Iris Flower Data Set 38

5.7 Fuzzy Type-1 Classifier Result on Tic-Tac-Toe Endgame Data Set 39

5.8 Fuzzy Type-2 Classifier Result on Breast Cancer Dataset 41

5.9 Fuzzy Type-2 Classifier Result on Iris Flower Data Set 43

5.10 Fuzzy Type-2 Classifier Result on Tic-Tac-Toe Endgame Data Set 44

5.11 Comparisons of Fuzzy Type-1 and Type-2 Classifiers Performance 46

5.12 Type-1 Classifier Implementation for Breast Cancer Data Set 47

5.13 Type-2 Classifier Implementation for Breast Cancer Data Set 55

6 CONCLUSIONS AND FUTURE WORK 58

6.1 Future Work 58

REFERENCES 59

LIST OF TABLES

Table 5.1: Description of the Selected Datasets……………………………………32

Table 5.2: Wisconsin Breast Cancer Dataset Attributes…………………….……33

Table 5.3: Descriptions of the Attributes ………………………………..……34

Table 5.4: Descriptions of Classes………………………………………………..34

Table 5.5: Classification Results for Fuzzy Type-1 on Breast Cancer Dataset.…….36

Table 5.6: Classification Results for Fuzzy Type-1 on Iris Flower Dataset……..…..38

Table 5.7: Classification Results for Fuzzy Type-1 on Tic-Tac-Toe Dataset………39

Table 5.8: Classification Results for Fuzzy Type-2 on Breast Cancer Dataset…......41

Table 5.9: Classification Results for Fuzzy Type-2 on Iris Flower Dataset……….43

Table 5.10: Classification Results for Fuzzy Type-2 on Tic-Tac-Toe Dataset……..44

Table 5.11: Classification Results for Type-1 Fuzzy Classifier…………………….46

Table 5.12: Classification Results for Type-2 Fuzzy Classifier ……………………46

Table 5.13: Train Set Samples………………………………………………….…...47

Table 5.14: Test Set Samples……………………………………………….......48

Table 5.15: Universes of Discourse Matrix Values…………………..…………..49

Table 5.16: Normalized Train Set Samples…………………………………………50

Table 5.17: Matrix (CC) of Cluster Centroid………………………………………..51

Table 5.18: Part of Matrix (MU) Degree of Memberships………………………...51

Table 5.19: Part of Matrix (RB) Rule Base……………………………………...53

Table 5.20: Defuzzification of Output Matrix. ……………………………………54

Table 5.21: Denormalized Output Matrix. ……………………………………..54

Table 5.22: Results Matrix……………………………………………………..54

Table 5.23: Matrix Rule Base. ……………………………………………….56

Table 5.24: Defuzzification Output Matrix. ………………………….……………56

Table 5.25: Denormalize Output Matrix ………………………….………..……57

Table 5.26: Results Matrix……………………………………...………………57

LIST OF FIGURES

Figure 3.1: The Common Membership Function of Fuzzy Sets [10]……………….10

Figure 3.2: Membership Function Terminology [11]...………………………....11

Figure 3.3: Three Dimensional Footprint of Uncertainty for The Type-2 Gaussian Membership Function [9]…………………………………………………...14

Figure 3.4: The Process of Union for Type-2 Gaussian Membership Function [9].15

Figure 3.5: Intersection for the Type-2 Gaussian Membership Functions [9]……..16

Figure 3.6: Block Diagram of Fuzzy Inference System [14]………………………..17

Figure 3.7: Block Diagram for a Mamdani System [14]………………..………….18

Figure 4.1: General Block Diagram of a Fuzzy Classifier…… …………………....22

Figure 4.2: The Fuzzy Type-1 Classifier Algorithm Flow Chart ………………......28

Figure 4.3: The Fuzzy Type-2 Classifier Algorithm Flowchart………...…………..30

Figure 5.1: The Membership Function of Type-1 Fuzzy Classifier on Breast Cancer Dataset………………………………………………………………………….37

Figure 5.2: The Membership Function of The Type-1 Fuzzy Classifier on Iris Flower Dataset. …………………………………………………………………….……..38

Figure 5.3: The Membership Function of Type-1 Fuzzy Classifier on Tic-Tac-Toe Dataset ……………………………………………..…………………….40

Figure 5.4: The Membership Function of Type-2 Fuzzy Classifier on Breast Cancer Dataset ……………………………...……………………………………………….42

Figure 5.5: The Membership Function of Type-2 Fuzzy Classifier on Iris Flower Dataset………………………………………………………………………….43

Figure 5.6: The Membership Function of Type-2 Fuzzy Classifier on Tic-Tac-Toe Dataset………………………………………………………………………….45

LIST OF ABBREVIATIONS

FCM	Fuzzy C-Mean Clustering.
FIS	Fuzzy Inference System.
FL	Fuzzy Logic.
FLS	Fuzzy Logic System
FOU	Footprint of Uncertainty.
KB	Knowledge Base.
MF	Membership Functions.
T1FS	Type-1 Fuzzy Set.
T2FS	Type-2 Fuzzy Set.

Chapter 1

INTRODUCTION

1.1 Fuzzy Logic – an Overview

Professor Lotfi Zadeh had emerged the notion of the fuzzy logic while improving the fuzzy set theory in 1965. Instead of constant and accurate reasoning, it provides approximate reasoning. In the fuzzy logic, there are different values rather than fixed values. There is two-valued logic in the binary sets which are true or false. On the other hand, the fuzzy logic contains various values (truth-value) ranging from zero to one. This notion of the fuzzy logic has been expanded for the fundamentals of limited certainty .

1.2 Fuzzy Model

A fuzzy model is a mathematics method. In the system of classification, rule-base of the fuzzy inferences had been frequently used. Inside this system, the relations within parameters appear via the if–then rules with vague predicates, such as if the intensity of the warmth is large then the heat will rise rapidly. In such cases, principle determines in a somewhat specified method in relation with the heat and the warmth inside the space when this model is activated. The sense for the large and rapid terminology will be described further accurately which had been completed via applying the concepts of fuzzy sets. The membership functions are described such that the items of the believed set have membership values in the range 0 to one. The value of one indicates complete membership and zero shows non-membership. The degree between the one and zero indicates partial membership in the set [1].

1.3 Fuzzy Classification

One of soft computing applications is classification which consists of assigning an object described by a set of features to a class. The object x $\in$ X is described by the vector of features v $\in$ V. Therefore it can equate object x as class membership with its feature values v $\bar{v}= [\bar{v}_1, \bar{v}_2 \; . \; . \; ., \bar{v}_n]$. Consequently, they can be used interchangeably x or $\bar{v}$. . It is assumed that fuzzy set A $\subseteq$ V is given as its membership function $\mu_{A(x)}=\mu_A(\bar{v}_1, \bar{v}_2 \; . \; . \; ., \bar{v}_n)$ $\bar{v}_i \in V_i$ for i= 1, . . . ,n. The fuzzy classifiers have been widely used according to their capability for utilizing the knowledge in the format of clear IF-THEN fuzzy rules. Within a classification task, rules are generally in the following form:

R^r: IF $v_1$is A_1^r AND v_2 is A_2^r. .. . . AND v_n is A_n^r, (1.1)

THEN $x \in w_1(\bar{z}_1^r)$, $x \in w_2$ $(\bar{z}_2^r)$, . . ., $x \in w_m$ $(\bar{z}_m^r)$

where r = 1, . . ., N, N is the number of rules and $\bar{z}_j^r$ is the membership degree of the object x to the *j*–the class according to rule [2].

1.4 Fuzzy C-Means Clustering

One of the widely used data clustering approaches is the fuzzy C-means (FCM). In this approach, each data point belongs to a cluster which is indicated by a membership grade. This method had been originally presented by the professor Jim Bezdek in 1981 for the improvement of the previous clustering technique. It provides a technique of how to group data points which populate some multidimensional range into a particular number of various clusters. One of the main benefits of the fuzzy c–means clustering algorithm is that it permits gradual membership of data points to clusters calculated as degrees in [0,1]. This provides a flexibility to explain that the data points can classify to more than one cluster [3].

1.5 Problem Statement

The data analysis for exploration is a crucial set of techniques applied in many areas such as science, engineering and business applications. The large datasets classification poses a problem practically in the area of machine learning datasets. Fuzzy classifiers are significant tools in this developing field. The scientists apply fuzzy rules and do not need regular assumptions to statistical classification. Though type-1 fuzzy logic classifier has been most applied to learning machine database intensively, it fails to tackle all kind of uncertainties, especially when more robust data are required. Type-2 fuzzy sets, an extension of type-1, simplifies the problem where the membership function is itself fuzzy with a three-dimensional representation. This is due to the property that type-2 fuzzy sets can shape uncertainties over membership functions and provide more degrees of freedom on their basic parameters.

1.6 Significance of Research

Fuzzy classes are suitable for the continuous data which does not fall certainly to discrete classes. As example, we can give plants type, soil classification, climatic data as well as many other geological, and medical and engineering applications. The Fuzzy classes can characterize the transitional areas more accurately than hard classifiers since the class membership is not binary (true/false) but instead, one location can belong to a few classes. Attribute ambiguity happens when class membership is vague or partial. Ambiguity is mainly a crisis for some remotely-sensed data. Fuzzy classes model the spatial and feature ambiguity nearby in most data sets more precisely than hard classifiers.

1.7 Thesis Organization

This study contains six sections. The first section highlights the basic introduction to fuzzy logic theory, fuzzy model and fuzzy classification. The second section of this study demonstrates the literature review for machine learning dataset based on fuzzy logic approaches. Chapter three discusses the fundamentals of fuzzy logic type-1 and type-2 models. The fourth section defines the methodology of the fuzzy classifiers. Chapter five presents the result of the simulation and outcome which are rules developed for the classification of machine learning dataset. Finally, chapter six includes a conclusion and future work.

Chapter 2

LITERATURE REVIEW OF THE RELEVANT RESEARCH

Classification and the clustering had been within the most significant works in a data mining, data analysis, and machine learning. Whereas, clustering is able to be seen as the most common form for unsupervised leanings and classification. The vagueness is been raised from a lack of knowledge, from a variety of classes of randomness and also from the disability to carry out sufficient dimensions [4].

Malek et al. suggested a model of automated detection, segmentation, and classification of breast cancer nucleus using a fuzzy system[5]. They first segmented the medical image by using an active shape for tracking and separating the cell of the nucleus in medical images. After that, they used nucleus that have been extracted for various textural features by applying the wavelet transforms to distinguish image using its texture. As a result, malign texture can be differentiated from benign one with the assumption that tumor texture is different from the texture of other kinds of tumors. At last, the researchers used the obtained features to be introduced as the input vector of a fuzzy C-means (FCM) clustering algorithm to categorize the images into malign and benign ones.

Yilmaz et al. analyzed the risk of breast cancer by implementing the fuzzy system. An earlier diagnosed cancer has been applied with this system and as a result, the cure was more efficient and in this theoretical method the threat of having breast tumor of the patient was reduced and the chance of preventing this threat would be proposed to people who have this syndrome. The patients stress of being affected by breast cancer will be tested after analyzing the threat rate of the tumor which is being assessed on the source of the self struggle of the individual to tumor expectancy of threat effect and aptitude to anxiety. To solve the problem, the existing images are being evaluated, sent to the system and the model has been given jointly by the fuzzy logic system like a new method. After getting high-quality results from the study, their method will create a pre-diagnosis for the individuals who probably have a threat of receiving tumor by the cause of working factors or lifestyle. As a result, this would help the individuals to obtain a kind of protection against the risk of tumor. Furthermore, the role of the fuzzy logic inference in the area of health and topics of artificial intelligence is tested in this method. Through this form of the method, patients will be able to obtain procedures against having a tumor and the average of having tumor might be reduced [6].

Indira et al. presented a new method for finding the breast cancer in women. Namely, they are Fuzzy c-means (FCM) algorithm and pattern recognition method. Algorithm had been used for the breast cancer clinic instances acquired from the University of Wisconsin. FCM algorithm utilized instances which were collected into two clusters, one with benign instances and other with malign instances. In addition, the input data has been divided into train data and test data and the achieved success of each was assessed. In pattern recognition technique every input data was allocated to one of the clusters acquired from the process of FCM classification [7].

Chen et al. introduced a new technique for managing classification problems based on the new fuzzy information measurement. Initially, they suggested the new fuzzy information need to measure a characteristic with respect to a set of training samples. Then, based on the suggested fuzzy information measurement, they have suggested an algorithm for building membership functions, calculating the class degree of each subset of training instances regarding to each class and calculating the fuzzy entropy of each subset of training instances. They have also suggested an estimation function for classifying testing instances [8].

Chapter 3

FUNDAMENTALS OF TYPE-1 AND TYPE-2 FUZZY LOGIC

The main idea of the fuzzy logic system is similar to human being's sense and the suggestion process. In the result of the fuzzy logic system, there is a derivative for the fuzzifications of joint inputs and outcomes by means of the related membership functions. The input of crisp system has been transformed into the dissimilar elements for the related membership function relying on its information. The main idea of the common fuzzy and the fuzzy logic are frequently utilized in the human being's routine in daily life. The computer systems are able to recognize each of the zeros or ones, as LOW or HIGH. This information is known as crisp information and is also able to be handled by every digital mechanism. The computer systems can handle the vague data with the help of human beings. The computers and machines handle those ambiguous data by using various fuzzy logic methods along with the knowledge of the fuzzy model.

3.1 Fuzzy Set Theory

The fundamental rules of the fuzzy sets concept differ from that of the traditional crisp sets generally in the degree to where a member belongs to a set. Within the crisp set approach, the members of the crisp set would not be members except if the membership is completely in that set and their membership is allocated a value of one. In contrast, the membership grade in the fuzzy set notion the set elements are defined in a way to allow a slow conversion from being a member of a set to a nonmember. Every variable has a grade of membership ranging from zero to one, where zero indicates nonmembership and one signify complete membership. The set of fuzzy approach is represented at the same time as a set of couples, *[x,* μ_A(x)*],* that x is a member in the universe of discourse *U*, and μ_A(x) is the degree of membership of x to the set A$\subset U$ .The universe of discourse *U* is discrete and fuzzy set *A* in this universe is referred via:

$$A=\sum_{x\in U} \mu_A(x)/x \qquad (3.1)$$

Within the fuzzy approach, fuzzy set A of universe *U* is described by function μ_A(x) known as the membership function of set *A*

$$\mu_A(x): U \to [0, 1], \text{ where } \mu_A(x) = 1 \text{ if } x \text{ is totally in } A; \qquad (3.2)$$

$$\mu_A(x) = 0 \text{ if } x \text{ is not in } A;$$

$$0 < \mu_A(x) < 1 \text{ if } x \text{ is partly in } A.$$

3.2 Properties of Fuzzy Sets

The height**:** This property of fuzzy a set A refers to the maximum grade of the membership function on *U* and it is defined as:

$$h(A) = \sup_{x \in U} \mu A(x). \qquad (3.3)$$

If *h(A)= 1*, *A* is called normal, else it is called subnormal.

Normality**:** A fuzzy set is said to be normal if at least one membership function of the set is equal to one.

Normalization**:** The normalization operation typically alters the scale of the membership function so the maximum value of the membership function is equal to *1*.

Fuzzy singleton: A fuzzy set is said to be a singleton if its support has a unique point $x \in U$ *with* μ_A(x) =1.

Convexity of Fuzzy Sets:

A fuzzy set *A* is convex if for any λ in *[0,1]*,

$$\mu_A(\lambda x1 + (1-\lambda)\, x2) \geq \mathrm{Min}\,(\mu_A(x1), \mu_A(x2)). \tag{3.4}$$

3.3 Membership Function of a Fuzzy Set

The membership function of a fuzzy set characterizes the degree of truth as an extension of estimation. The Membership functions of the fuzzy sets are able to be defined in any numeral ways. The form of the membership function utilized determines the fuzzy set and so the decision on which type to utilize depends on the purpose. The membership function selection is a subjective characteristic for the fuzzy logic; it permits the preferred rates to be explained correctly. The main membership functions widely used applications are mentioned below [9].

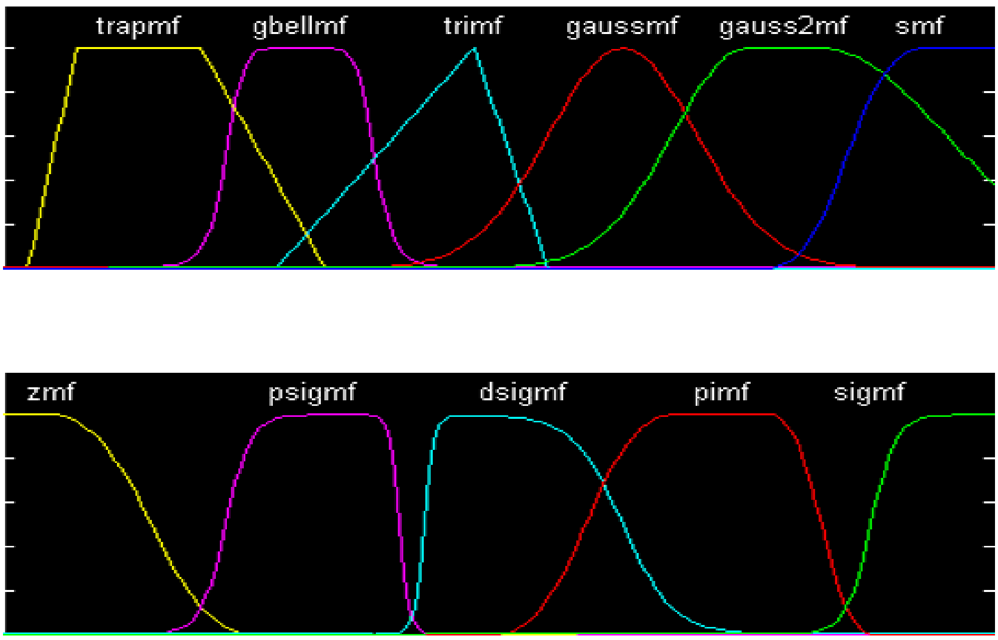


Figure 3.1: The Common Membership Function of Fuzzy Sets [10].

3.3.1 Membership Function Terminology

Core: The core of a set A is the crisp subset of A with membership grade equal to one. So,

$$\text{Core }(A) = \{x|\mu_A(x) = 1\} \quad [9]. \tag{3.5}$$

Alpha-cut: An α-cut or α-level set of a fuzzy set $A \subseteq X$ is an ordinary set $A\alpha \subseteq X$, such that:

$$A\alpha=\{\mu_A(x)\geq\alpha, x\in X\} \quad [9]. \tag{3.6}$$

Support: The Support of a fuzzy set A is the crisp set of all points in the Universe of Discourse U such that the membership function of A is non-zero.

$$\text{supp}(A) = \{x|\mu A(x) > 0\} \quad [9]. \tag{3.7}$$

Crossover point: The Crossover point of a fuzzy set is the element in U at which its membership function value is 0.5 [9].

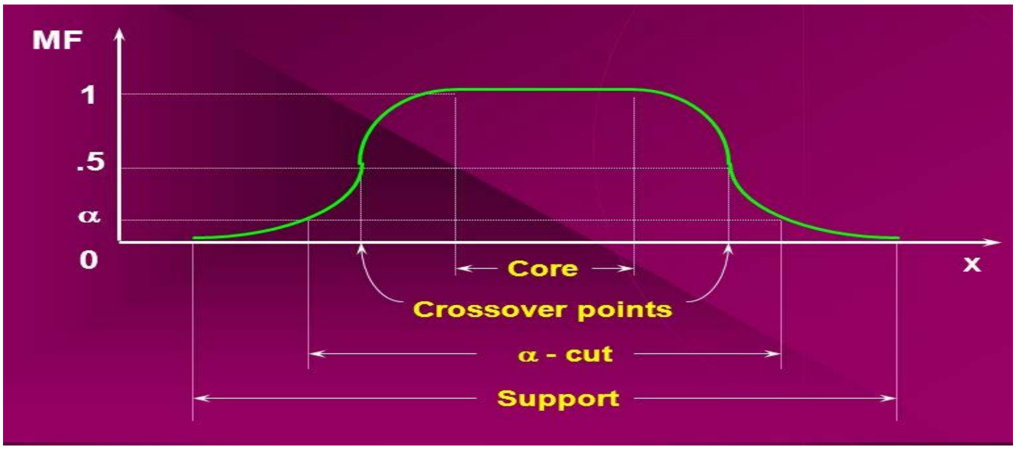


Figure3.2: Membership Function Terminology [11].

3.4 Type-2 Fuzzy Sets and Systems

The type-2 fuzzy logic methods are rule-based fuzzy methods which are compatible to type-1 fuzzy logic methods within the concept of the structure. Within the type-2 fuzzy sets, there are two major shapes of type-2 fuzzy sets: interval and general type-2 fuzzy sets. In the practice, the interval shape is the most utilized shape in type-2 fuzzy sets due to its simplicity of calculations compared to the generalized membership shape. Type-2 fuzzy logic methods usually need extra calculations than type-1 fuzzy logic methods, however, they present more elasticities and freedom in symbolizing knowledge [12].

3.4.1 Fuzzy Sets and Uncertainty

The presence of uncertainties and the require of information in numerous of real world problems makes it complex to model some real world problems. Uncertainty in fuzzy set theory has been perceived and split into three kinds which are:

- Fuzziness (or vagueness): This results from the imprecise boundaries of fuzzy sets.
- Non-specificity (or imprecision): It is linked to relevant sets of alternatives when some alternative belongs to a specific set of alternatives but they do not know which one in the set it is.
- Strife (or discord): which expresses conflicts among the various sets of alternatives. [13]

The type-1 fuzzy logic had been utilized effectively within a large set of problems. However type-1 methods are not capable of modeling ambiguity in complex environments. Problems associated with modeling uncertainty utilizing crisp membership functions of type-1 fuzzy sets had been perceived by many researchers. Type-2 fuzzy logic systems may encompass a lot of benefits compared with type-1 fuzzy logic systems. These benefits comprise of:

- The type-2 fuzzy set is capable of handling the numerical and linguistic vagueness in membership function and comprises a footprint of uncertainty (FOU).
- Type-2 fuzzy sets symbolize the inputs and the outputs using fewer principles.
- The Type-2 fuzzy sets entrenched a huge number of type-1 fuzzy sets to explain elements through a detailed explanation.
- The additional dimension supplied via the footprint of uncertainty (FOU) permits the type-2 fuzzy logic system to create outputs that cannot be accomplished via the type-1 fuzzy logic system utilizing the similar number of membership functions [12].

Two circumstances have to be believed concerning the common understanding which the general type-2 fuzzy logic system should outperform the interval from which also should do better than type-1 fuzzy logic system. These two circumstances are the reliance of acting on the selection of the form parameters in addition to on the unevenness of vagueness in the implementation. Consequently, a superior selection of the model's elements utilizing automated techniques is preferred to gain additional clearer outcomes concerning this comparison [12].

3.4.2 Type-2 Fuzzy Sets

Fuzzy type-2 set stands for $\widetilde{A}$ is described by means of a type-2 membership function $\mu_{\tilde{A}}(x, \mu)$ which x $\in X$ and $\mu \in \mu_x \subseteq [0; 1]$: such that :

$$\tilde{A} = \{ ((x;\mu), \mu_{\tilde{A}}((x;\mu)) | \; \forall x \in X; \; \forall u \in \mu_x \subseteq [0; 1] \} \qquad (3.8)$$

$\tilde{A}$ is also called an interval fuzzy type-2 set. Interval fuzzy type-2 sets are simpler to calculate general fuzzy type-2 sets..In Figure 3.3 the footprint of uncertainty for a type-2 Gaussian membership function is presented.

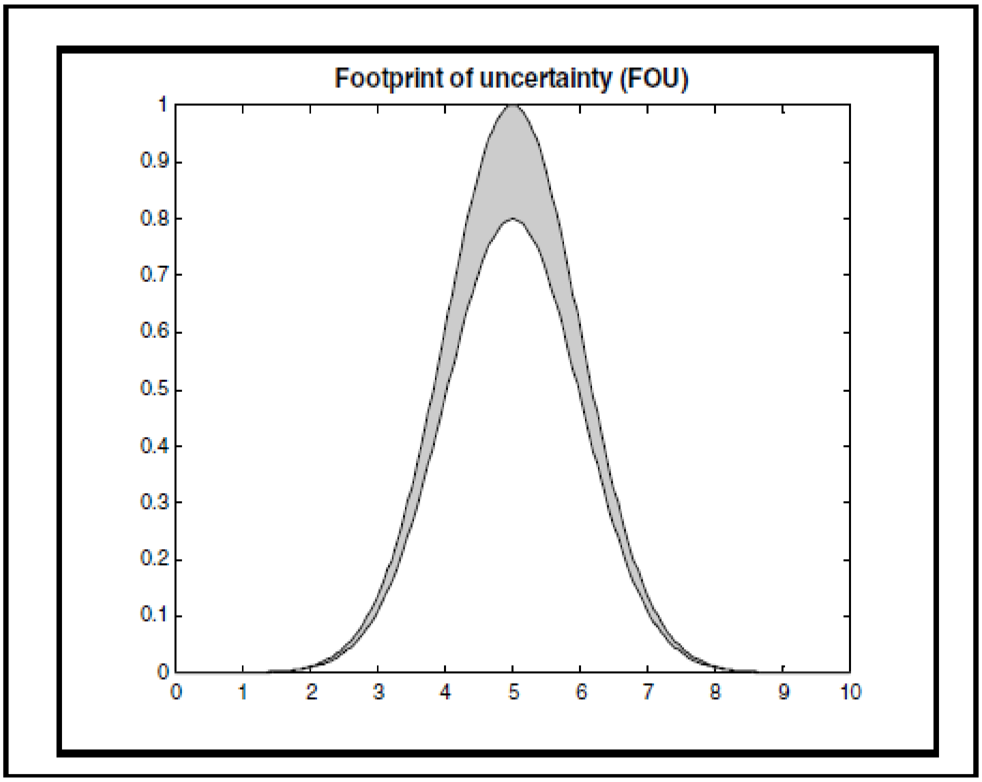


Figure3.3: Three Dimensional Footprint of Uncertainty for The Type-2 Gaussian Membership Function [9].

An upper membership function and lower membership functions are two type-1 membership functions which are boundaries for the footprint of uncertainty (FOU) of the type-2 fuzzy set Ã. The upper membership function is related to the upper boundary of FOU (Ã). The lower membership function is associated with the lower boundary of FOU(Ã) [9].

3.4.3 Operations of Type-2 Fuzzy Sets

The operation of the type-2 fuzzy sets has been described. A mechanism of calculating the union, intersection, also complement for type-2 fuzzy sets is Zadeh extension principle. Consider type-2 fuzzy sets $\tilde{C}1$ and $\tilde{C}2$,

Union for Type-2 Fuzzy Sets:

The union of $\tilde{C}1$ and $\tilde{C}2$ is a new type-2 fuzzy set, similar to a union of type-1 fuzzy sets *C1* and *C2*. Union of two type-2 fuzzy sets is illustrated in Figure 3.4.

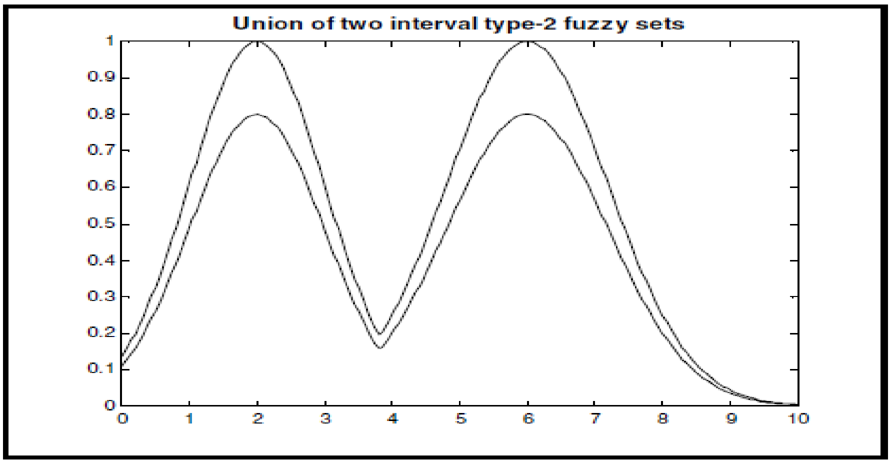


Figure 3.4: The Process of Union For Type-2 Gaussian Membership Functions [9].

Intersection of Type-2 Fuzzy Sets:

The intersection of $\tilde{C}1$ and $\tilde{C}2$ is another type-2 fuzzy set, similar to the intersection of type I fuzzy sets *C1* and C2. The intersection of two type-2 Gaussian MFs is demonstrated in Figure3.5.

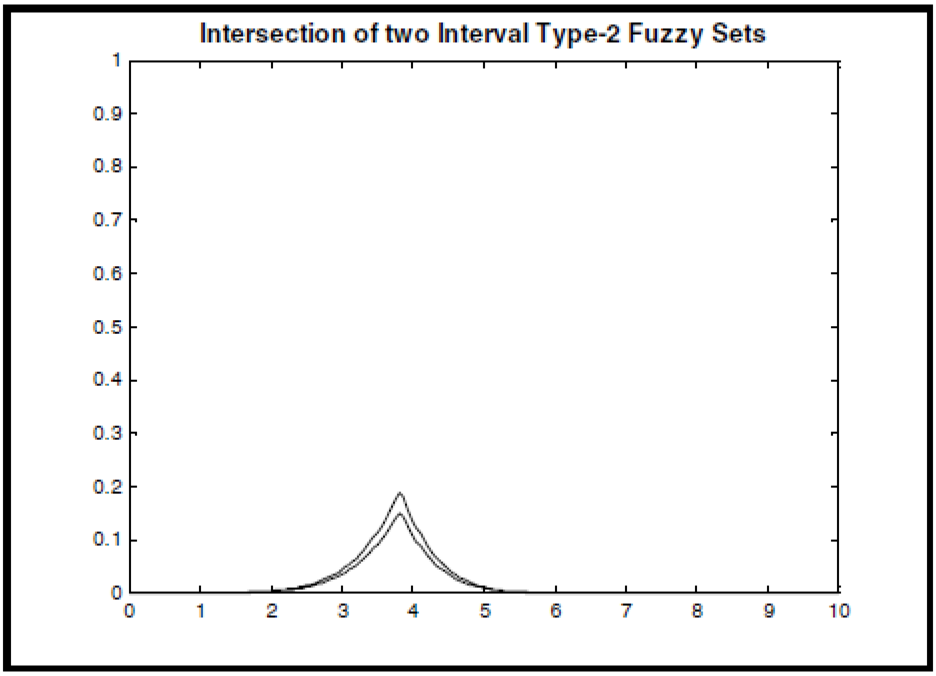


Figure 3.5: Intersection for The Two Type-2 Gaussian Membership Functions [9].

Complement of a Type-2 Fuzzy Set

Complement of set $\tilde{C}1$ is a different type-2 fuzzy set; it is only at a same as a complement for type I fuzzy set *C*. It definition is

$$\mu_{\tilde{C}}\tilde{1}(x) = 1-\mu_{\tilde{C}}\tilde{1}(x). \qquad (3.9)$$

3.5 Fuzzy Inference System

Fuzzy inference is the operation of creating the charting from the given input to an output utilizing the fuzzy logic. The charting offers a substructure from which decisions are able to be complete. The operation of fuzzy inference comprises all of the parts that are explained in Membership Functions, Logical Operations, and If-Then Rules. A popular form of MFs is triangular, even though trapezoidal, Gaussian and bell curvatures are also utilized, however, the contour is commonly a lesser amount of significant than the numeral of curvatures and their positioning. Starting from 3 to 7 curvatures are normally suitable for covering a needed set of the entered value or the universe of discourse within fuzzy terminology. As explained previously, the inference procedure depends on the combination of logic rules in the form of IF THEN statements, which the IF section is known the antecedent and the THEN section is known as the consequent The Fuzzy inference methods had been effectively applied in many areas, for example, automatic control, data classification, decision analysis, expert systems, and computer vision for the reason that of its multidisciplinary nature, the fuzzy inference system is recognized by the numerous of names, for instance, fuzzy logic controller, fuzzy rule-based system, fuzzy expert system, fuzzy model, fuzzy associative memory, and simply fuzzy system [9].

3.5.1 The Architecture of Fuzzy Inference Systems

Block diagram of a fuzzy system is given in Figure 3.6.

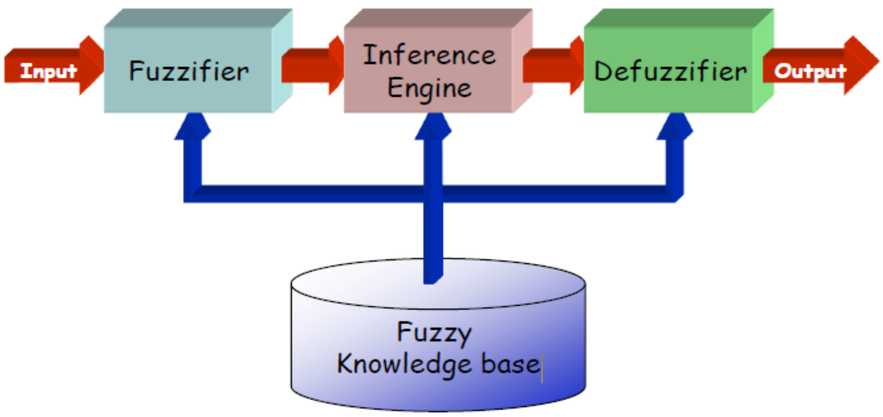


Figure 3.6: Block Diagram of Fuzzy Inference System [14].

Fuzzy Knowledge Base:

The rule-base and the database are together known as the knowledge base.

- The rule -base consists of the numeral value of fuzzy IF–THEN rules.
- The database that describes the membership functions of the fuzzy sets utilized in the fuzzy rules [14].

Fuzzification:

The initial phase to implement is to transmit the input parameters of a fuzzy system to membership functions (MF) that have a variety of forms. The simple membership functions are formed by using Gaussian distribution curve, triangular and trapezoidal membership functions. The grade of the membership function is computed by means of putting the selected input variable on the horizontal axis, while the vertical axis illustrates quantification of the degree of membership of the input variable. The input parameter will obtain a single membership function and a weighting factor will be used with the values of each input parameter. The grade of the influence or degree of membership (DOM) has been indicated via these weighting factors. Consequently, by calculating the logical product of the membership weights for every active rule, the output response for the fuzzy set measures are created. [9].

Inference Engine :

Using IF-Then type fuzzy rules transfers the fuzzy input to the fuzzy output, as shown in Figure 3.7.

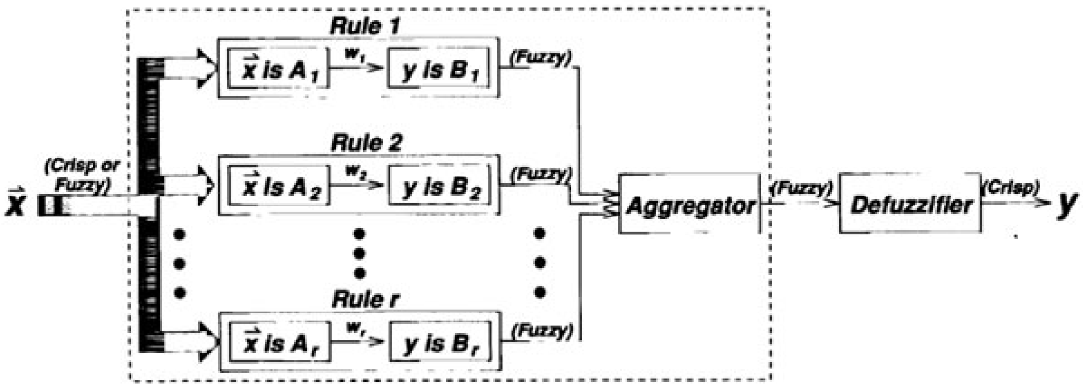


Figure 3.7: Block Diagram for Mamdani System [14].

Defuzzification:

It is the procedure that generates an experimental outcome within traditional logic by giving the fuzzy sets and the matching membership grades. The defuzzification operation is explaining the membership grades of the fuzzy sets to a particular choice or valid information, where the easiest although minimum practical defuzzification technique is to select the set by means of the maximum membership. A problem with this method is to miss data. The popular also practical defuzzification method is the center of gravity. The primary step for this technique is the fuzzy centroid .The x coordinate of the centroid is the defuzzified value. Five commonly used defuzzifying methods are Centroid of the area (COA), Bisector of the area (BOA), Mean of maximum (MOM), Smallest of maximum (SOM) and Largest of maximum (LOM) [15].

3.5.2 Fuzzy Rule Based System

The popular two significant kinds are Mamdani and Sugeno fuzzy rule based systems. The Mamdani type is the most popular fuzzy rule-based system. An additional famous fuzzy rule based system that is known as Sugeno or Takagi–Sugeno Kang technique of fuzzy controller. The major dissimilarity of these two techniques is found inside the consequent of fuzzy rules [15].

3.5.3 Mamdani Fuzzy Rule Based System

Mamdani's fuzzy rule-based system is generally the well-known fuzzy inference method. This approach was among initial control systems created utilizing fuzzy set theory. Mamdani's attempt had been depending on Professor Lotfi Zadeh's principles on fuzzy algorithms for complex systems and decision processes [15].

- The steps of Mamdani inference procedure

  1. The Fuzzification of input parameters ,

2. Implement the fuzzy operation,
3. Implement implication method,
4. Implement aggregation method ,
5. Defuzzification.

Chapter 4

IMPLEMENTATION OF THE PROPOSED SYSTEM

4.1 Introduction

This chapter describes first, the general structure and the flowchart of proposed fuzzy classifier. Then, the algorithm and flowchart of the type-1 fuzzy classifier will be explained. Furthermore, type-2 fuzzy classifier algorithm and flowchart will be defined. Finally, the methods which were used for evaluating the fuzzy classification will be described.

4.2 Methodology

In this thesis fuzzy logic type-1 and type-2 were used for the classification of the machine learning data sets. The classification is done by applying two types of fuzzy sets: fuzzy type-1 algorithm uses Mamdani type fuzzy inference and Gaussian membership functions. Also, the comparisons between two fuzzy set classifiers were applied and the result showed that type-2 fuzzy classifier algorithm outperforms type-1 fuzzy logic. Block diagram of fuzzy classifiers is shown in Figure 4-1.

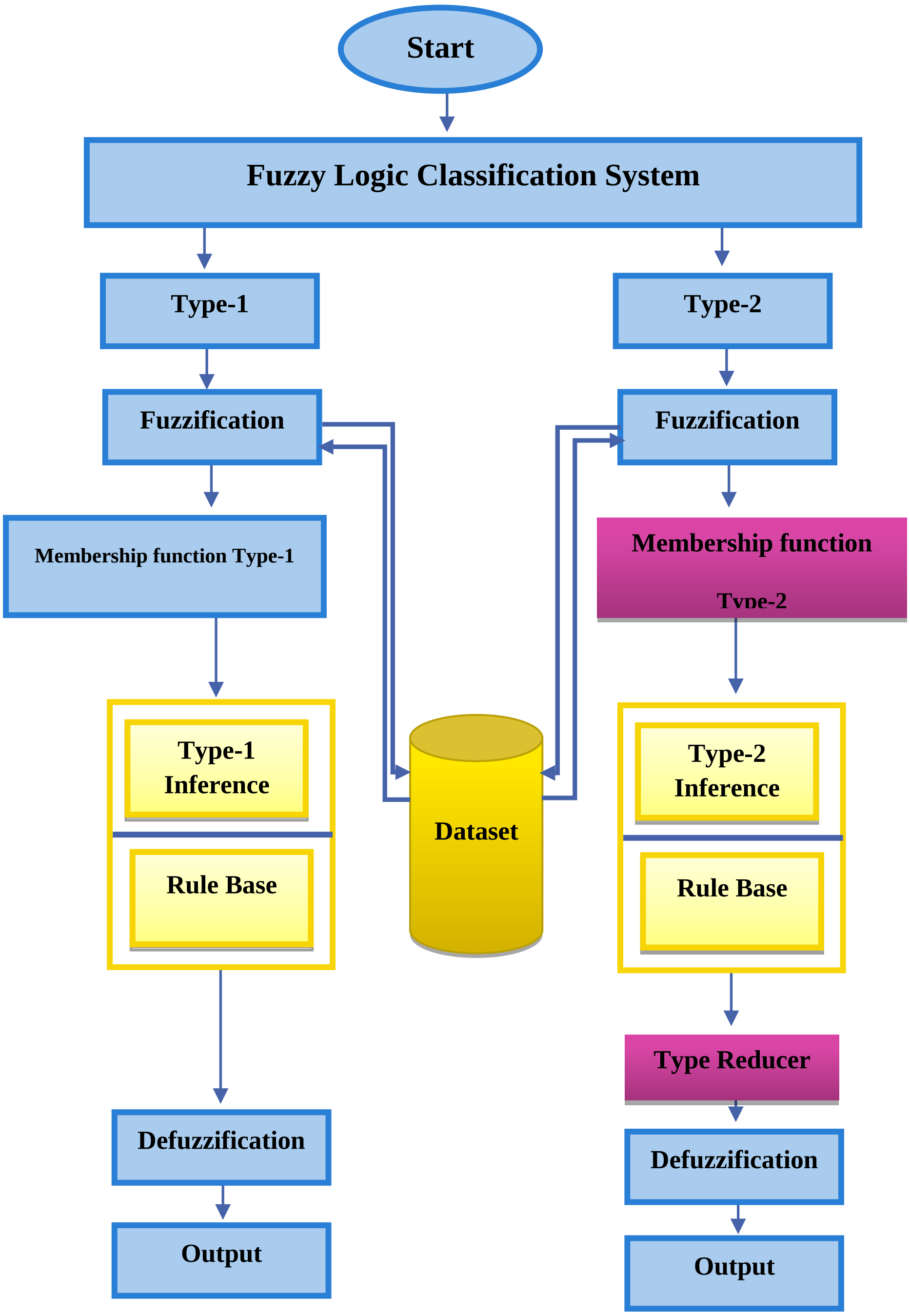


Figure 4.1: Block Diagram of a Fuzzy Classifier.

4.3 Principles of Fuzzy Classifier

4.3.1 Type-1 Classifier General Process Steps

Type-1 fuzzy classifier process steps consist of:

1-Fuzzification: Fuzzify input variables by converting the crisp numerical values of input variables into the corresponding membership values of the suitable fuzzy sets by using type1 Gaussian membership function.

2-Type-1 Inference Engine:

a) Apply Fuzzy Operator

When the fuzzy inference system consists of more than one input variable, the antecedent of IF-Then rule can usually be determined via fuzzy linguistic sets. Since in most cases every input variable has one corresponding fuzzy set, the degree of membership is dependent on which to be defined. The most popular fuzzy operations are OR operation and the AND operation and to define these logical operators the function max and function min are implemented.

b) Apply Implication Method

The consequent part of IF-Then rule is to figure out an additional fuzzy linguistic set via a suitable membership function but not like the outcome from antecedent part of IF-Then rule that a single numerical value is produced. The inference technique in IF-Then part is to reshape the fuzzy set of consequent part according to the result related with the antecedent, or say the single number. This procedure is called implication method. The AND operation is applied which truncates the fuzzy set of consequent part.

c) Apply Aggregation Method

The subsequent to each of the IF-Then rules is producing a modified fuzzy set as output, the aggregation process is applied to join these fuzzy sets that symbolize the outputs of rules into a particular fuzzy set in order to find the decision. The last joining fuzzy set is the output of the aggregation procedure, and all output variable of the fuzzy inference system will have a single corresponding combined fuzzy set for reference. Function max, sum, and probabilistic OR are all applicable for aggregation operation.

3-Defuzzification

The defuzzification is the opposite process of fuzzification. As in the first process, this is the process of extracting an accurate quantity out of the domain of fuzzy set to the output variable. The Centroid Method is used in this research.

4.3.2 Type-2 Classifier Procedural Steps

A type-2 fuzzy logic system is a rule-based system that is similar to a type-1 fuzzy logic system in terms of its structure and components. The only difference is that a type-2 fuzzy logic system uses at least one type-2 fuzzy set and it has an extra output process component which is called the type-reducer before defuzzification as shown in figure 4.1. The type-reducer reduces output type-2 fuzzy sets to type-1 fuzzy sets then the defuzzifier reduces it to a crisp output. The components of a type-2 Mamdani fuzzy system are [14] :

1-Fuzzification: Fuzzify input variables by converting the crisp numerical values of input variables into the corresponding membership values of the suitable fuzzy sets by using type-2 Gaussian membership function which is shown in figure 4.5.

2-Rules

A fuzzy rule in type-2 fuzzy logic system is also a conditional statement in the shape of IF-THEN where it includes two elements, the IF part known the antecedent part and the THEN part known as the consequent part. The difference from type-1 fuzzy logic system rules is that, rules in type-2 fuzzy logic system utilize type-2 fuzzy sets. Utilizing one type-2 fuzzy set is adequate to determine the system as a type-2 fuzzy logic system [14].

3-Inference Engine

Inference Engine chart input type-2 fuzzy sets into output type-2 fuzzy sets via stratifying the consequent part where this procedure of charting from the antecedent part into the consequent part is interpreted as a type-2 fuzzy implication which needs calculations of union and intersection of type-2 fuzzy sets. The inference engine in a Mamdani type-2 fuzzy logic system charts the input type-2 fuzzy sets into the output type-2 fuzzy sets. The rules in Mamdani model have type-2 fuzzy sets in both the antecedent and the consequent parts.

4-Output Processor

There are two stages in the output process:

a) Type-Reduction: main rule of the Type-reducer reduces type-2 fuzzy sets that have been generated by the inference engine to type-1 fuzzy sets by performing a centroid calculation (Mendel, 2001). Type-Reducer replaces each type-2 consequent set by its centroid which is a type-1 set and then calculates a weighted average of these centroids to get a type-1 fuzzy set. This is the bottleneck of the type-2 fuzzy logic systems as this process requires expensive computations especially when using non-interval type-2 sets [14].

b) Defuzzification: Defuzzifier maps the reduced output type-1 fuzzy sets that have been reduced by the type-reducer into crisp values exactly as the case of defuzzification in type-1 fuzzy logic systems. Any defuzzification methods of type-1 fuzzy sets can be used here [14].

4.4 Fuzzy Type-1 Classifier Algorithm

The steps of fuzzy type-1 are as follows:

Step 1: Load the Selected dataset and randomly divide it to two set (train set, test set).

Step 2: Determine the structure of the type-1 fuzzy set, the number of rows and columns.

Step 3: Determine the universes of the input and output variables.

Step 4: Normalize input data

Step 5: Determine the fuzzy rules base using clustering. Fuzzy c-means clustering was used to group the dataset to determine the feasibility of a fuzzy rule base.

Step 6: Determining the membership functions for the individual variables. Gaussian Type-1 membership function was used.

Step 7: Determine and optimize the fuzzy inference engine using Mamdani's methods: Max-Min Inference.

***Step 8**:* Determine the output data to non-fuzzy values using defuzzification Centroid Method.

Step 9: Denormalize the fuzzy set.

Step 10: Evaluate the type-1 classifier.

Flowchart of a fuzzy type-1 classier is given in Figure 4.2.

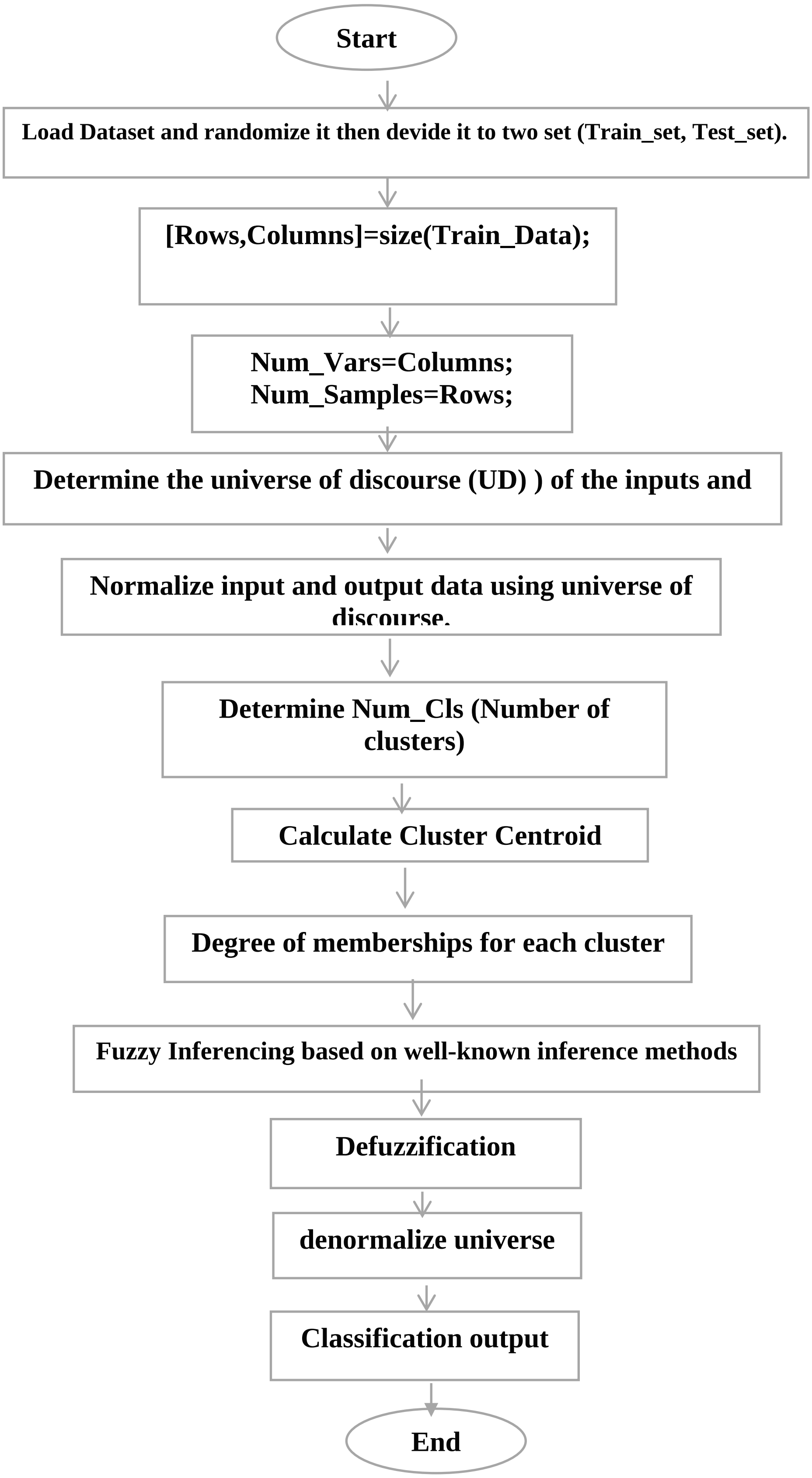


Figure 4.2: The Fuzzy Type-1 Classifier Algorithm Flow Chart.

4.5 Fuzzy Type-2 Classifier Algorithm

The steps of fuzzy type-2 algorithm are as follows:

Step 1: Load the Selected Dataset Data and randomize it then divide it to two set (train set, test set).

Step 2: Determine the structure of the interval type-2 fuzzy set, the number of row and Column.

Step 3: Determine the universes of the input and output variables.

Step 4: Normalize input data

Step 5: Determine the fuzzy rules base using clustering. Fuzzy c-means clustering was used to group the dataset to determine the feasibility of a fuzzy rule base.

Step 6: Determining the membership functions for the individual variables. Gaussian Type-2 Membership function, was used.

Step 7: Determine and optimize the fuzzy inference engine using Mamdani's Methods: Max-Min Inference.

Step 8: Convert output data from type-2 to type-1 using type reducer method.

Step 9: Determine the type-1 output data to non-fuzzy values using defuzzification Centroid method.

Step 10: Denormalize the fuzzy set.

Step 11: Evaluate the type-2 classifier.

Flowchart of a fuzzy type-2 classifer is given in Figure 4.3.

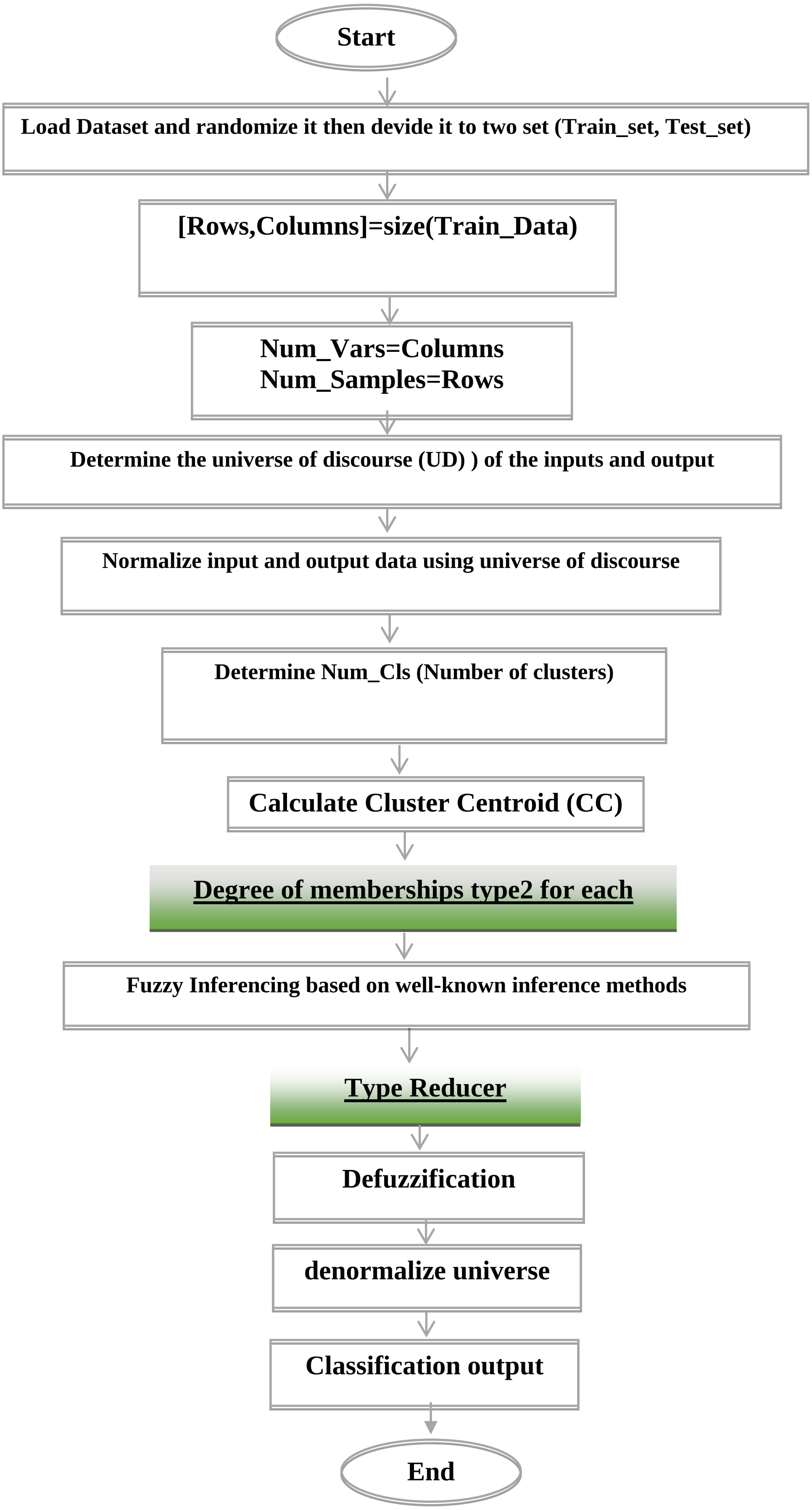


Figure 4.3:The Fuzzy Type-2 Classifier Algorithm Flow Chart.

4.6 Evaluate The Fuzzy Classification

The method were used to measure the performance of the proposed fuzzy classifier is The Total Accuracy which is defined as follows:

$$Acc = \frac{\text{Number of Correct Classification instances}}{\text{Total Number of Training Data instances}}$$

Chapter 5

RESULTS AND DISCUSSION

5.1 Experiments and Results

Three datasets were used to compare type-1 fuzzy and ype-2 fuzzy logic systems. The data sets that have been selected are Wisconsin breast cancer dataset, iris fisher dataset and tic-tac-toe dataset.. Table (5-1) gives a brief description of the datasets: note that they exhibit a wide variety of characteristics [16].

Table 5.1: Description of the Selected Datasets

Dataset Name	Sample Size	Number of attributes	Number of class
Wisconsin Breast Cancer	699	9	2
Iris	150	4	3
Tic-tac-toe	958	9	2

MATLAB (2015a) programming platform and a personal computer with a CPU: T6500, RAM: 3.00GB, Operating System: Windows7 are used to conduct The experiments. Details of datasets used in experimental work are given in The following sections.

5.2 Wisconsin Breast Cancer Data Set

The details of the attributes found in this dataset are listed in table 5.2.

Table 5.2: Wisconsin Breast Cancer Dataset Attributes

	Attribute	Domain
1	Sample code number	**ID number**
2	Clump Thickness	1 - 10
3	Uniformity of Cell Size	1 - 10
4	Uniformity of Cell Shape	1 - 10
5	Marginal Adhesion	1 – 10
6	Single Epithelial Cell Size	1 – 10
7	Bare Nuclei	1 – 10
8	Bland Chromatin	1 – 10
9	Normal Nucleoli	1 – 10
10	Mitoses	1 - 10
11	Class	2 for benign, 4 for malignant

Brief explanations of these features are as follows.In the Clump thickness, benign cells tend to be grouped in monolayers, while cancerous cells are often grouped in multilayer. While in the Uniformity of cell size/shape the cancer cells tend to vary in size and shape. That is why these parameters are valuable in determining whether the cells are cancerous or not. In the case of Marginal adhesion, the normal cells tend to stick together, where cancer cells tend to lose this ability. So loss of adhesion is a sign of malignancy. In the Single epithelial cell size the size is related to the uniformity mentioned above. Epithelial cells that are significantly enlarged may be a malignant cell. The Bare nuclei is a term used for nuclei that are not surrounded by cytoplasm (the rest of the cell). Those are typically seen in benign tumors. The Bland Chromatin describes a uniform "texture" of the nucleus seen in benign cells. In cancer cells, the chromatin tends to be coarser. The Normal nucleoli are small structures seen in the nucleus. In normal cells, the nucleolus is usually very small if visible. In cancer cells, the nucleoli become more prominent, and sometimes there are more of them. Finally, Mitoses is nuclear division plus cytokines and produce two identical daughter cells during prophase. It is the process in which the cell divides and replicates. Pathologists can determine the grade of cancer by counting the number of mitoses.

5.3 Iris Flower Data Set Description

As a reference dataset to show some simple concepts in multivariate statistics, the Iris dataset collected by Anderson (1936) is used here as well. It is most popular due to the publication by Fisher (1936). The data originates from three different species of iris flowers, Iris setosa, Iris versicolor, and Iris virginica. For each of the species a sample of 50 flowers was examined by measuring the length and the width of the sepals and petals. The dataset is of special interest since it constitutes a difficult separation problem. There are in total 4 numerical attributes and no missing value in the dataset as described in Tables 5.3 and 5.4.

Table 5.3: Descriptions of The Attributes

Attribute	Description
A1	Sepal length in cm
A2	Sepal width in cm
A3	Petal length in cm
A4	Petal width in cm

Table 5.4: Descriptions of Classes

Class	Description
C1	Iris Setosa
C2	Iris Versicolour
C3	Iris Virginica

5.4 Tic-Tac-Toe Endgame Data Set

This database encodes the complete set of possible board configurations at the end of tic-tac-toe games, where "x" is assumed to have played first. The target concept is "win for x" (i.e., true when "x" has one of 8 possible ways to create a "three-in-a-row"). Number of Instances is 958 (legal tic-tac-toe endgame boards) an number of Attributes is 9, each corresponding to one tic-tac-toe square Attribute information symbols are named as follows : x=player, x has taken, o=player, o has taken, b=blank.

Attribute Information of this data set can be listed as follows:

1. top-left-square: {x,o,b}
2. top-middle-square: {x,o,b}
3. top-right-square: {x,o,b}
4. middle-left-square: {x,o,b}
5. middle-middle-square: {x,o,b}
6. middle-right-square: {x,o,b}
7. bottom-left-square: {x,o,b}
8. bottom-middle-square: {x,o,b}
9. bottom-right-square: {x,o,b}
10. Class: {positive,negative}.

5.5 Fuzzy Type-1 Classifier Result on Breast Cancer Dataset

Table 5.5 shows the classification results for the breast cancer dataset which have been selected after using fuzzy type-1 classifier. The accuracy for breast cancer dataset is 92.97, with true positives= 321 and false negatives = 29.

Table 5.5: Classification Results for Fuzzy Type-1 on Breast Cancer Dataset

#	Breast cancer
Accuracy	92.97
Number of training samples	350
Number of test samples	350
True positives	321
False negatives	29

Membership functions for features and class variables for The Wisconsin breast cancer dataset are shown in Figure 5.1.

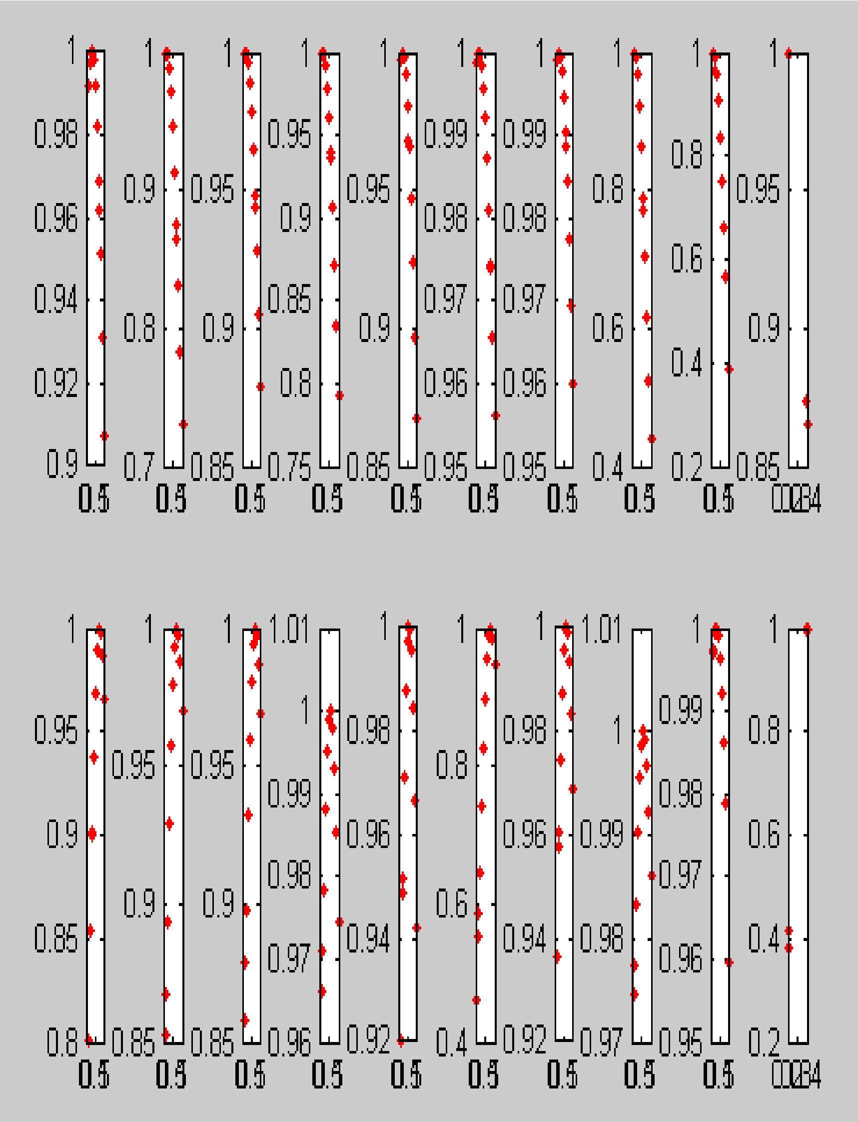


Figure 5.1: The Membership Functions of The Type-1 Fuzzy Classifier on Breast Cancer Dataset.

5.6 Fuzzy Type-1 Classifier Result on Iris Flower Data Set

Table 5.6 shows the classification results for the Iris flower data set that have been selected after using Fuzzy Type-1 classifier. The accuracy for Iris flower data set is 88.08 with true positives= 66 and false negatives =9.

Table 5.6: Classification Results for Fuzzy Type-1 on Iris Flower Dataset

#	Iris flower
Accuracy	88.08
Number of training samples	75
Number of test samples	75
True positives	66
False negatives	9

Membership Functions for features and class variables for Iris dataset are given in Figure 5.2.

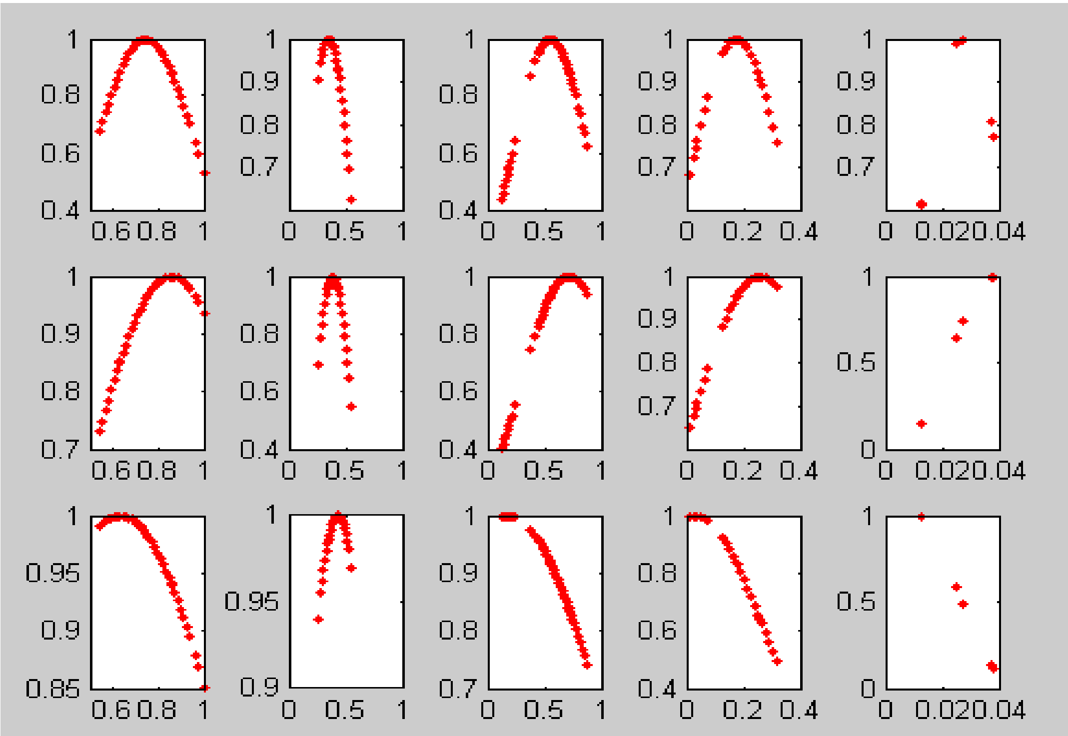


Figure 5.2:The Membership Function of Type-1 Fuzzy Classifier on Iris Flower Dataset.

5.7 Fuzzy Type-1 Classifier Result on Tic-Tac-Toe Endgame Data Set

Table 5.7 shows the classification results for the Tic-Tac-Toe data set that have been selected after using Fuzzy Type-1 classifier. The accuracy is 90.8,with true positives= 435 and false negatives = 44.

Table 5.7: Classification Results for Fuzzy Type-1 on Tic-Tac-Toe Dataset

#	**Tic-Tac-Toe**
Accuracy	90.8
Number of training samples	479
Number of test samples	479
True positives	435
False negatives	44

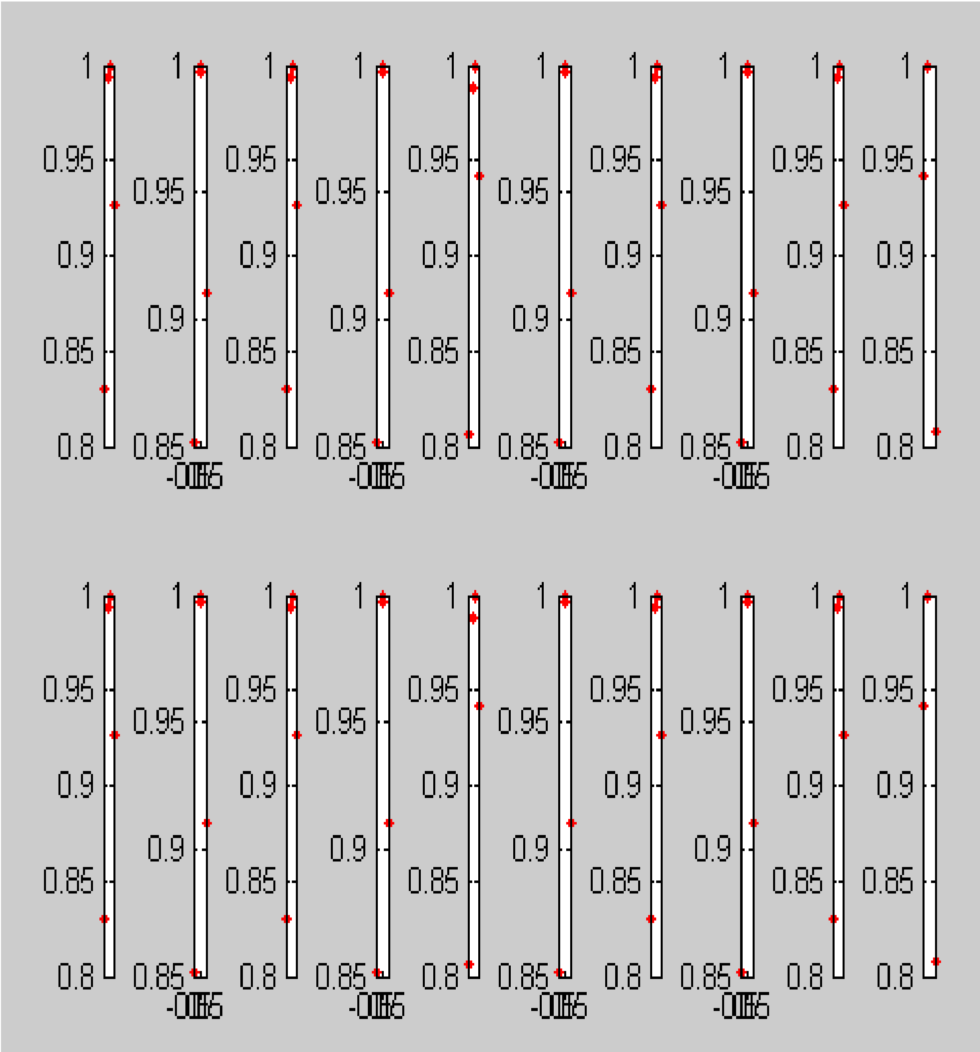


Figure 5.3 : The Membership Function of Type-1 Fuzzy Classifier on Tic-Tac-Toe Dataset

5.8 Fuzzy Type-2 Classifier Result on Breast Cancer Dataset

Table 5.8 shows the classification results for the breast cancer data set that have been selected after using Fuzzy Type-2 classifier. The accuracy is 97.9, with true positives = 342 and false negatives = 8.

Table 5.8: Classification Results for Fuzzy Type-2 on Breast Cancer Dataset

#	**Breast cancer**
Accuracy	97.9
Number of training samples	350
Number of test samples	350
True positives	342
False negatives	8

Fuzzy type-2 membership functions for features and class variables for Wisconsin breast cancer dataset is shown in Figure 5.4.

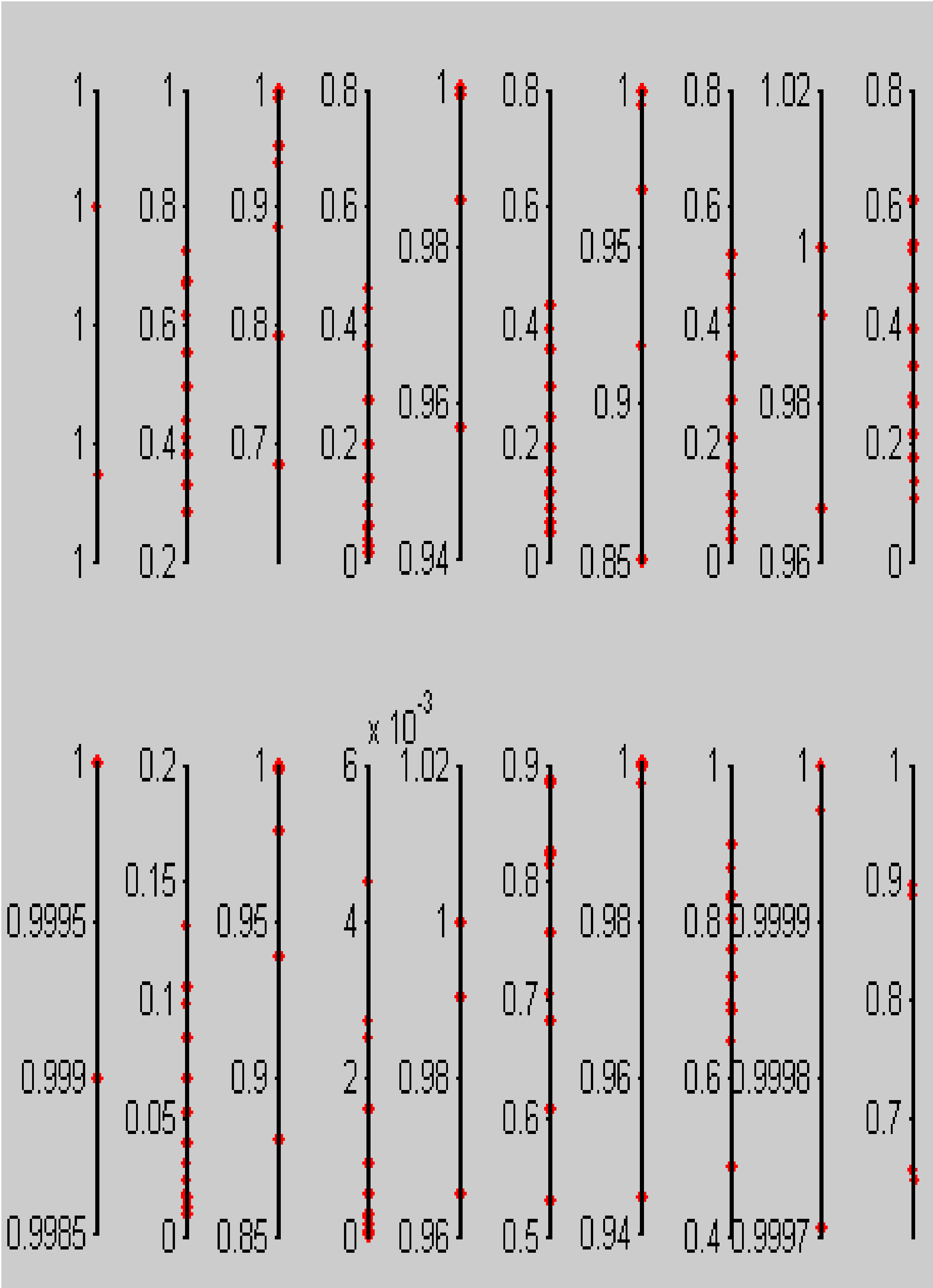


Figure 5.4 :The Membership Function of Type-2 Fuzzy Classifier on Breast Cancer Dataset.

5.9 Fuzzy Type-2 Classifier Result on Iris Flower Data Set

Table 5.9 shows the classification results for the Iris flower data set that have been selected after using Fuzzy Type-2 classifier. The accuracy for Iris flower data set is 96, with true positives =72 and false negatives= 3.

Table 5.9: Classification Results for Fuzzy Type-2 on Iris Flower Dataset

#	Iris flower
Accuracy	96
Number of training samples	75
Number of test samples	75
True positives	72
False negatives	3

Fuzzy type-2 membership functions for feature and class variables for The Iris flower dataset is given in Figure 5.5.

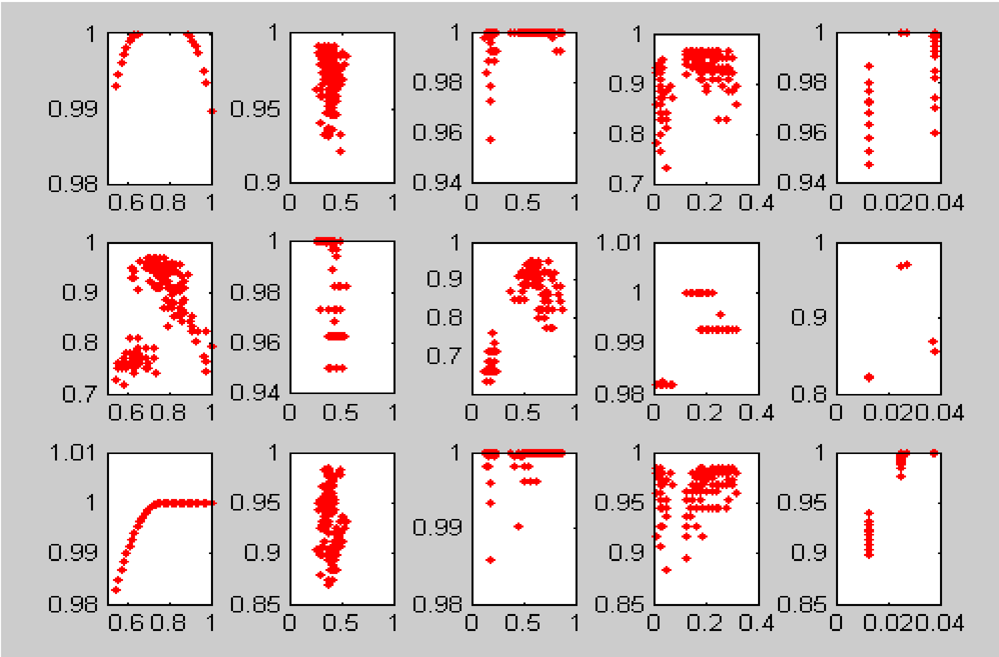


Figure 5.5: The Membership Function of The Type-2 Fuzzy Classifier on Iris Flower Dataset.

5.10 Fuzzy Type-2 Classifier Result on Tic-Tac-Toe Endgame Data Set

Table 5.10 shows the Classification Results for the Tic-Tac-Toe data set that have been selected after using Fuzzy Type-2 classifier. The RMSE for Tic-Tac-Toe data set is 0.223840, the Acc 94.9, true positive 455 and false negative was 24.

Table 5.10: Classification Results for Fuzzy Type-2 on Tic-Tac-Toe Data Set

#	**Tic-Tac-Toe**
Accuracy	94.9
Number of training samples	479
Number of test samples	479
True positives	455
False negatives	24

Fuzzy type-2 membership functions for feature and class variables for The Tic-tac-toe dataset are illustrated in Figure 5.6.

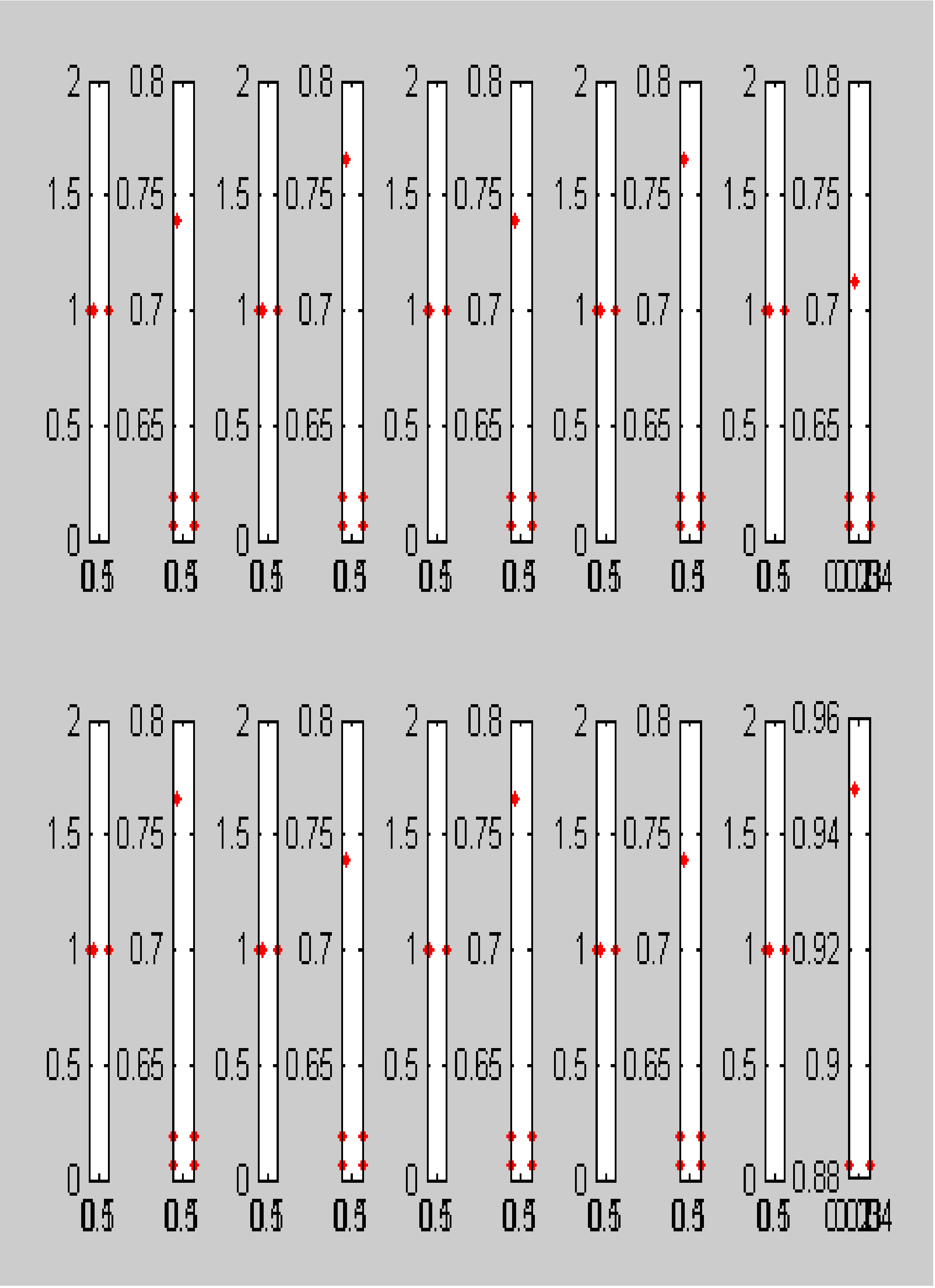


Figure 5.6: The Membership Function of The Type-2 Fuzzy Classifier on Tic-Tac-Toe Dataset.

5.11 Comparisons of Fuzzy Type-1 and Type-2 Classifiers Performance

After calculating the accuracy, True positives and False negatives, the two classifiers were compared to conclude on Their comparative the best classifier. Tables (5-11, 5-12), give the comparison results with the two classifiers. It is seen type-2 classifier is generally better than fuzzy type-1 classifier.

Table 5.11: Classification Results for Type-1 Fuzzy Classifier

#	Breast Cancer	Iris Dataset	tic-tac-toe
Accuracy	92.97	88.03	90.8
Number of training samples	350	75	479
Number of test samples	350	75	479
True positives	321	66	435
False negatives	29	9	44

Table 5.12: Classification Results for Type-2 Fuzzy Classifier

#	Breast Cancer	Iris Dataset	Ti350c-tac-toe
Accuracy	97.9	96	94.9
Number of training samples	350	75	479
Number of test samples	350	75	479
True positives	342	72	455
False negatives	8	3	24

5.12 Type-1 Classifier Implementation for Breast Cancer Data Set

Step 1: Load the Selected Dataset Data and divide it to two sets as train set 25 samples and test set 25 samples.

Table 5.13:Train Set Samples

Train Set Samples (25 samples)										
ID	**Clump Thickness**	**Uniformity of Cell**	**Uniformity of Shape**	**Marginal Adhesion**	**Single Epithelial Cell Size**	**Bare Nuclei**	**Bland Chromatin**	**Normal Nucleoli**	**Mitoses**	**Class**
557583	5	10	10	10	10	10	10	1	1	4
1202253	5	1	1	1	2	1	1	1	1	2
1174131	10	10	10	2	10	10	5	3	3	4
1058849	5	1	1	1	2	1	1	1	1	2
1165926	9	6	9	2	10	6	2	9	10	4
1200847	6	10	10	10	8	10	10	10	7	4
770066	5	2	2	2	2	1	2	2	1	2
1257366	3	1	1	1	2	1	1	1	1	2
1313982	4	3	1	1	2	1	4	8	1	2
1168278	3	1	1	1	2	1	2	1	1	2
897471	4	8	8	5	4	5	10	4	1	4
601265	10	4	4	6	2	10	2	3	1	4
128059	1	1	1	1	2	5	5	1	1	2
486283	3	1	1	1	2	1	3	1	1	2
695091	5	10	10	5	4	5	4	4	1	4
1071760	2	1	1	1	2	1	3	1	1	2
1290203	3	1	1	1	2	1	2	1	1	2
877291	6	10	10	10	10	10	8	10	10	4
1234554	1	1	1	1	2	1	2	1	1	2
1091262	2	5	3	3	6	7	7	5	1	4
1054593	10	5	5	3	6	7	7	10	1	4
1267898	5	1	3	1	2	1	1	1	1	2
255644	10	5	8	10	3	10	5	1	3	4
640744	10	10	10	7	9	10	7	10	10	4
1324572	5	1	1	1	2	1	2	2	1	2

Table 5.14: Test Set Samples

Test Set Samples (25 samples)										
ID	**Clump Thickness**	**Uniformity of Cell**	**Uniformity of Shape**	**Marginal Adhesion**	**Single Epithelial Cell Size**	**Bare Nuclei**	**Bland Chromatin**	**Normal Nucleoli**	**Mitoses**	**Class**
1333877	5	4	5	1	8	1	3	6	1	2
488173	1	4	3	10	4	10	5	6	1	4
625201	8	2	1	1	5	1	1	1	1	2
1116116	9	10	10	1	10	8	3	3	1	4
1243256	10	4	3	2	3	10	5	3	2	4
1213375	8	4	4	5	4	7	7	8	2	2
1213383	5	1	1	4	2	1	3	1	1	2
1184184	1	1	1	1	2	5	1	1	1	2
566346	3	1	1	1	2	1	2	3	1	2
871549	5	1	2	1	2	1	2	1	1	2
1276091	1	3	1	1	2	1	2	2	1	2
688033	1	1	1	1	2	1	1	1	1	2
1145420	6	1	1	1	2	1	2	1	1	2
1017023	6	3	3	5	3	10	3	5	3	2
653777	8	3	4	9	3	10	3	3	1	4
1306282	6	6	7	10	3	10	8	10	2	4
869828	1	1	1	1	1	1	3	1	1	2
752904	10	1	1	1	2	10	5	4	1	4
1266154	8	7	8	2	4	2	5	10	1	4
822829	8	10	10	10	6	10	10	10	10	4
1168359	8	2	3	1	6	3	7	1	1	4
1043068	3	1	1	1	2	1	2	1	1	2
1157734	4	1	1	1	2	1	3	1	1	2
1238021	1	1	1	1	2	1	2	1	1	2
1239347	8	7	8	5	10	10	7	2	1	4

Step 2: Determine the structure of the type-1 fuzzy set, the number of row and Column.

Columns=11 , Rows=350

Step 3: Determine the universes of the input and output variables.

Min_Data=0 , Max_Data=10

Table 5.15: Universes of Discourse Matrix Values.

Universes of Discourse matrix values							
No	UD	No	UD	No	UD	No	UD
1	0.1	26	2.4	51	4.9	76	7.4
2	0	27	2.5	52	5	77	7.5
3	0.1	28	2.6	53	5.1	78	7.6
4	0.2	29	2.7	54	5.2	79	7.7
5	0.3	30	2.8	55	5.3	80	7.8
6	0.4	31	2.9	56	5.4	81	7.9
7	0.5	32	3	57	5.5	82	8
8	0.6	33	3.1	58	5.6	83	8.1
9	0.7	34	3.2	59	5.7	84	8.2
10	0.8	35	3.3	60	5.8	85	8.3
11	0.9	36	3.4	61	5.9	86	8.4
12	1	37	3.5	62	6	87	8.5
13	1.1	38	3.6	63	6.1	88	8.6
14	1.2	39	3.7	64	6.2	89	8.7
15	1.3	40	3.8	65	6.3	90	8.8
16	1.4	41	3.9	66	6.4	91	8.9
17	1.5	42	4	67	6.5	92	9
18	1.6	43	4.1	68	6.6	93	9.1
19	1.7	44	4.2	69	6.7	94	9.2
20	1.8	45	4.3	70	6.8	95	9.3
21	1.9	46	4.4	71	6.9	96	9.4
22	2	47	4.5	72	7	97	9.5
23	2.1	48	4.6	73	7.1	98	9.6
24	2.2	49	4.7	74	7.2	99	9.7
25	2.3	50	4.8	75	7.3	100	9.8

Table 5.16: Normalized Train Set Samples.

Normalized Train Set Samples (25 samples)										
ID	**Clump Thickness**	**Uniformity of Cell**	**Uniformity of Shape**	**Marginal Adhesion**	**Single Epithelial Cell Size**	**Bare Nuclei**	**Bland Chromatin**	**Normal Nucleoli**	**Mitoses**	**Class**
557583	0.5	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.1	0.5
1202253	1	1	1	0.2	1	1	0.5	0.3	0.3	1
1174131	0.5	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.1	0.5
1058849	0.9	0.6	0.9	0.2	1	0.6	0.2	0.9	1	0.9
1165926	0.6	1	1	1	0.8	1	1	1	0.7	0.6
1200847	0.5	0.2	0.2	0.2	0.2	0.1	0.2	0.2	0.1	0.5
770066	0.3	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.1	0.3
1257366	0.4	0.3	0.1	0.1	0.2	0.1	0.4	0.8	0.1	0.4
1313982	0.3	0.1	0.1	0.1	0.2	0.1	0.2	0.1	0.1	0.3
1168278	0.4	0.8	0.8	0.5	0.4	0.5	1	0.4	0.1	0.4
897471	1	0.4	0.4	0.6	0.2	1	0.2	0.3	0.1	1
601265	0.1	0.1	0.1	0.1	0.2	0.5	0.5	0.1	0.1	0.1
128059	0.3	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.1	0.3
486283	0.5	1	1	0.5	0.4	0.5	0.4	0.4	0.1	0.5
695091	0.2	0.1	0.1	0.1	0.2	0.1	0.3	0.1	0.1	0.2
1071760	0.3	0.1	0.1	0.1	0.2	0.1	0.2	0.1	0.1	0.3
1290203	0.6	1	1	1	1	1	0.8	1	1	0.6
877291	0.1	0.1	0.1	0.1	0.2	0.1	0.2	0.1	0.1	0.1
1234554	0.2	0.5	0.3	0.3	0.6	0.7	0.7	0.5	0.1	0.2
1091262	1	0.5	0.5	0.3	0.6	0.7	0.7	1	0.1	1
1054593	0.5	0.1	0.3	0.1	0.2	0.1	0.1	0.1	0.1	0.5
1267898	1	0.5	0.8	1	0.3	1	0.5	0.1	0.3	1
255644	1	1	1	0.7	0.9	1	0.7	1	1	1
640744	0.5	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.1	0.5
1324572	1	1	1	0.2	1	1	0.5	0.3	0.3	1

Step 5: Determine the fuzzy rules base using clustering. Fuzzy c-means clustering was used to group the dataset to determine the feasibility of a fuzzy rule base.

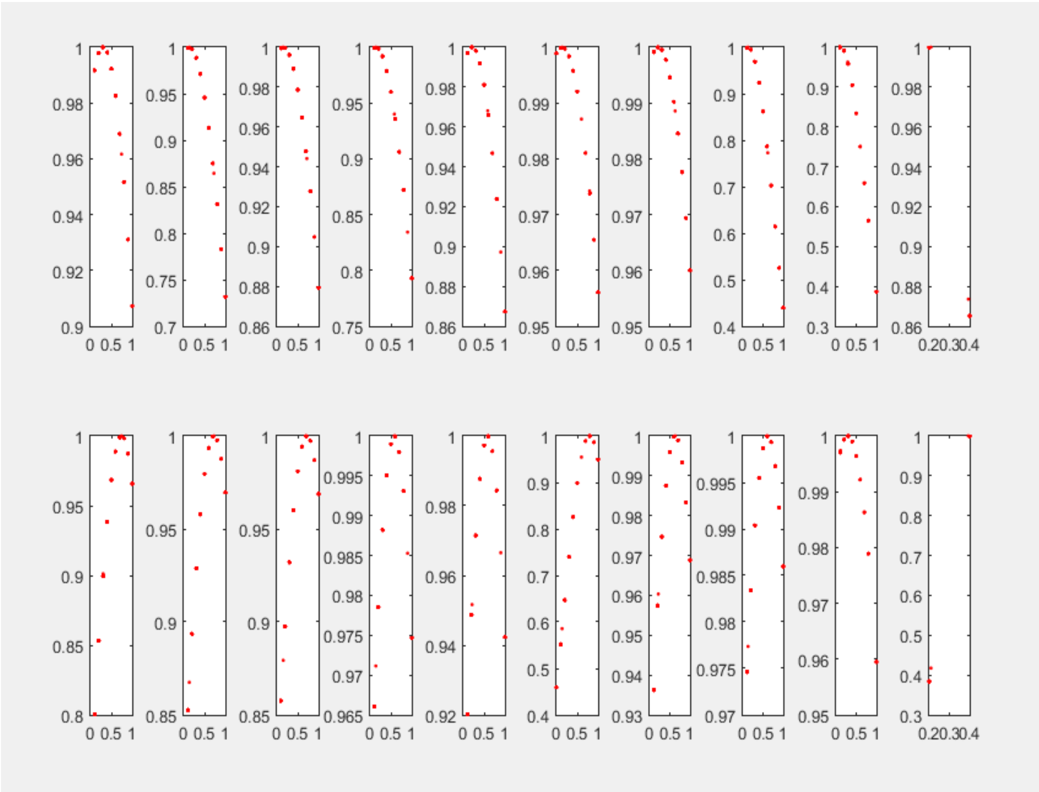
Table 5.17: Matrix (CC) of Cluster Centroid .

Matrix (CC) of Cluster Centroid									
0.3037	0.1348	0.1527	0.1384	0.2113	0.1354	0.215	0.1291	0.1123	0.2087
0.7448	0.7253	0.7191	0.5833	0.5879	0.7945	0.6316	0.6168	0.2903	0.3935

Table 5.18: Part of Matrix (MU) Degree of Memberships.

Part of Matrix (MU) Degree of memberships for each cluster									
0.1957	0.975	0.1657	0.975	0.2609	0.1469	0.9703	0.9915	0.7646	0.9966
0.8043	0.025	0.8343	0.025	0.7391	0.8531	0.0297	0.0085	0.2354	0.0034

Step 6: Determining the membership functions for the individual variables. Gaussian Type1 membership function.



Step 7: Determine and optimize the fuzzy inference engine.

Table 5.19: Part of Matrix (RB) Rule Base.

Part of Matrix (RB) Rule Base									
0.9923	0.7314	0.8793	0.7931	0.8674	0.9561	0.96	0.9991	0.9998	0.8651
0.9923	0.9995	0.9995	0.9995	1	0.9999	0.9991	0.9991	0.9998	0.9997
0.9072	0.7314	0.8793	0.9988	0.8674	0.9561	0.9946	0.9689	0.9584	0.8651
0.9923	0.9995	0.9995	0.9995	1	0.9999	0.9991	0.9991	0.9998	0.9997
0.9311	0.9135	0.9048	0.9988	0.8674	0.9871	1	0.5253	0.3865	0.8651
0.9825	0.7314	0.8793	0.7931	0.9238	0.9561	0.96	0.4397	0.6592	0.8651
0.9923	0.9982	0.9996	0.9988	1	0.9999	1	0.9946	0.9998	0.9997
1	0.9995	0.9995	0.9995	1	0.9999	0.9991	0.9991	0.9998	0.9997
0.9981	0.9887	0.9995	0.9995	1	0.9999	0.9977	0.6141	0.9998	0.9997
1	0.9995	0.9995	0.9995	1	0.9999	1	0.9991	0.9998	0.9997
0.9981	0.8312	0.9277	0.96	0.9919	0.992	0.96	0.9236	0.9998	0.8651
0.9072	0.971	0.9891	0.9356	1	0.9561	1	0.9689	0.9998	0.8651
0.9917	0.9995	0.9995	0.9995	1	0.992	0.9946	0.9991	0.9998	0.9997
1	0.9995	0.9995	0.9995	1	0.9999	0.9995	0.9991	0.9998	0.9997
0.9923	0.7314	0.8793	0.96	0.9919	0.992	0.9977	0.9236	0.9998	0.8651
0.9978	0.9995	0.9995	0.9995	1	0.9999	0.9995	0.9991	0.9998	0.9997
1	0.9995	0.9995	0.9995	1	0.9999	1	0.9991	0.9998	0.9997
0.9825	0.7314	0.8793	0.7931	0.8674	0.9561	0.9776	0.4397	0.3865	0.8651
0.9917	0.9995	0.9995	0.9995	1	0.9999	1	0.9991	0.9998	0.9997
0.9978	0.9458	0.9961	0.9919	0.966	0.981	0.9845	0.8616	0.9998	0.8651

Step 8: Determine the output data to non-fuzzy values using defuzzification Centroid Method.

Table 5.20 : Defuzzification of Output Matrix.

Part of defuzzification of output matrix									
2.7999	2.6217	2.7999	2.6217	3.1152	3.055	2.6217	2.6188	2.6819	2.6188

Step 9: Denormalize the fuzzy set.

Table 5.21: Denormalized Output Matrix.

Part of Denormalized output matrix									
0.28	0.262	0.28	0.262	0.312	0.306	0.262	0.262	0.268	0.262

Step 10: Identify the class.

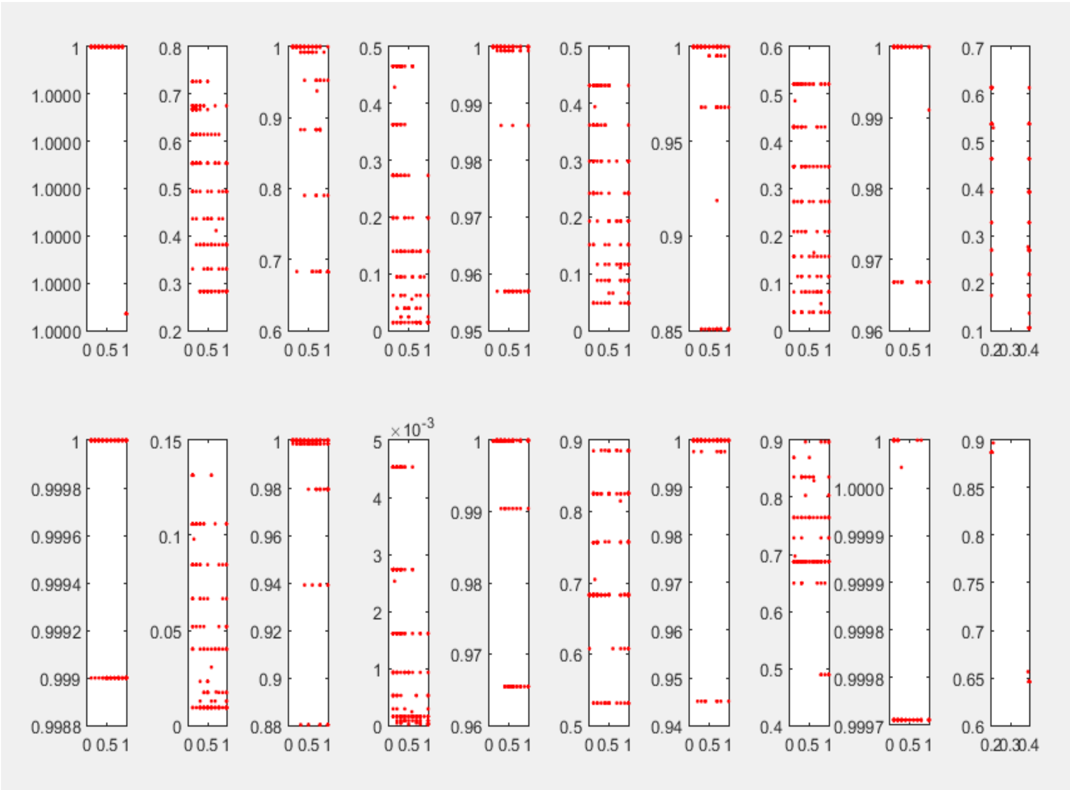
Table 5.22: Results Matrix.

Part of Results matrix									
2	4	2	4	2	2	4	4	4	4

5.13 Type-2 Classifier Implementation for Breast Cancer Data Set

Step 1 to 5 for type-2 fuzzy classifier are The same as Those of fuzzy type-1 classifier.

Step 6: Determining the membership functions for the individual variables. Gaussian Type-2 membership function.



Step 7: Determine and optimize the fuzzy inference engine.

Table 5.23: Matrix Rule Base.

Part of Matrix (RB) Rule Base									
1	1	1	1	1	1	1	1	1	1
0.55371	0.554	0.283	0.554	0.33	0.49391	0.554	0.675	0.615	0.675
0.68299	1	0.683	1	0.992	0.68299	1	1	1	1
0.01399	0.466	0.014	0.466	0.094	0.01399	0.363	0.466	0.273	0.466
0.9569	1	0.957	1	0.986	0.9569	1	1	1	1
0.04815	0.432	0.048	0.432	0.066	0.04815	0.362	0.432	0.432	0.432
0.85068	1	1	1	1	0.85068	1	1	1	1
0.03837	0.521	0.43	0.521	0.43	0.03837	0.43	0.521	0.521	0.521
0.96682	1	0.967	1	0.967	0.99997	1	1	1	1
0.10512	0.537	0.105	0.537	0.105	0.17384	0.537	0.537	0.537	0.537
0.999	1	0.999	1	1	0.999	1	1	1	1
0.0093	0.106	0.009	0.106	0.031	0.0093	0.106	0.106	0.106	0.106
0.88024	1	1	1	1	0.88024	1	1	1	1
2.11E-05	0.005	5E-04	0.005	0.003	2.11E-05	0.003	0.005	9E-04	0.003
0.9999	1	1	1	0.99	0.96543	1	1	1	1
0.68285	0.683	0.825	0.683	0.608	0.53123	0.757	0.683	0.684	0.683
0.99998	1	1	1	0.945	1	1	1	1	1
0.68708	0.687	0.836	0.687	0.489	0.72878	0.687	0.687	0.687	0.687
0.99971	1	1	1	1	0.99971	1	1	1	1
0.64634	0.888	0.646	0.888	0.646	0.64634	0.888	0.888	0.888	0.888

Step 8*:* Determine the output data to non-fuzzy values using defuzzification Centroid Method.

Table 5.24: Defuzzification Output Matrix.

Part of defuzzification output matrix									
0.268	0.221	0.268	0.221	0.268	0.268	0.235	0.221	0.244	0.221

Step 9: Denormalize the fuzzy set.

Table 5.25: Denormalize Output Matrix

Part of Denormalize output matrix									
2.07014	2.571	2.075	2.571	2.356	2.07014	2.553	2.571	2.535	2.553

Step 10: Identify the class.

Table 5.26: Results Matrix

Part of Results matrix									
2	4	2	4	2	2	4	4	4	4

Chapter 6

CONCLUSIONS AND FUTURE WORK

1. The type-2 fuzzy classifiers are able to perform better than type-1 fuzzy classifiers which have additional design parameters. Consequently, type-2 fuzzy classifiers are more attractive than the type-1 classifiers considering to interpretability and precision. The essential benefits of the type-2 fuzzy logic classifiers are their ability to manipulate more vagueness or in the condition of extra uncertainty.

2. The main significant section of the type-2 fuzzy logic control system appears to be the membership function about the source. Consequently, the simplified type-2 fuzzy logic controls the membership function close to the source .

6.1 Future Work

One can apply the genetic algorithm with the type-1 fuzzy set and type-2 fuzzy set for the optimization of membership parameters . Another suggestion is applying neural networks with type 1 fuzzy set and type 2 fuzzy set for classification.

REFERENCES

[1] Heinz, D., (2009, October 11). *Control systems robotics and automation* Book. isbn:184826156x,9781848261563.

[2] Scherer, R., (2012) .*Studies in fuzziness and soft computing*.Book. doi: 1434-9922.

[3] Bora, J., D., & Gupta, K., A., (2014,April). *A comparative study between fuzzy clustering algorithm and hard clustering algorithm*.International Journal of Computer Trends and Technology (IJCTT*)*. issn: 2231-2803.

[4] Forghani, Y., Yazdi, S., H., & Effati, S., (2012,September 7).*Classification of fuzzy data based on the support vector machines*.*Expert Systems Article*. doi: 10.1111/j.1468-0394.2012.00645.x.

[5] Malek, J., & Sebri, A., & Mabrouk, S., Torki, K., & Tourki, R., (2008,June12).*Automated breast cancer diagnosis based on gvf-snake segmentation, wavelet features extraction and fuzzy classification*. Journal of Signal Processing Systems , 55(13), 49-66.doi :10.1007/s11265-008-0198-2.

[6] Yilmaz, A., & Ayan, K., (2011). *Risk analysis in breast cancer disease by using fuzzy logic and effects of stress level on cancer risk*. Scientific Research and Essays,vol. 6(24), pp. 5179-5191.doi: 10.5897/sre11.840.

[7] Indira, M., (2013). *Fuzzy analysis of breast cancer disease using fuzzy c-means and pattern recognition*. Journal System ,vol 2,No 1.

[8] Chen, A., b., & Jen-Da Shie, A., (2009). *Fuzzy classification systems based on fuzzy information gain measures*. Expert Systems with Applications, 36(3), pp.4517-4522.

[9] Castillo, O., & Melin, P., (2008).*Type-2 fuzzy logic: theory and applications*. Book, vol. 223.isbn 978-3-540-76283-6.

[10] The Most Common Membership Functions.Retrieved from https://www.cavin.edu/~priberio/othrlnks/Fuzzy/fuzzyops.htm.

[11] Membership Function Terminology.Retrieved from http://images.slideplayer.co 27/9251317/slides/slide_10.jpg.

[12] Mendel, J., M., (2007). *Advances in type-2 fuzzy sets and systems*.Information sciences 177(1), 84-110,doi:10.1016/j.ins.2006.05.003.

[13] Klir, G., & Wierman, M., (1999).*Uncertainty-Based Information: Elements of Generalized Information Theory*.Book. DOI:10.1007/978-3-7908-1869-7.

[14] Shomeand, A., & Ashok, D., S., (2012,June*). Fuzzy logic approach for boiler temperature & water level control*. International Journal of Scientific & Engineering Research, vol 3, issue 6.issn 2229-5518.

[15] Castillo, O., & Melin, P., (2012). *Recent advances in interval type-2 fuzzy systems*. Book. doi: 10.1007/978-3-642-28956-9_1.

[16] Machine Learning Datasets. (2016).Retrieved from https://archive.ics.uci.edu/m1 /datasets.html.