

2007

Computer-aided weld inspection by fuzzy modeling with selected features

Sean Najm Ghazavi

Louisiana State University and Agricultural and Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_theses



Part of the [Construction Engineering and Management Commons](#)

Recommended Citation

Ghazavi, Sean Najm, "Computer-aided weld inspection by fuzzy modeling with selected features" (2007).
LSU Master's Theses. 452.

https://digitalcommons.lsu.edu/gradschool_theses/452

This Thesis is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Master's Theses by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.

COMPUTER-AIDED WELD INSPECTION BY FUZZY MODELING WITH SELECTED FEATURES

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
In partial fulfillment of the
Requirements for the degree of
Master of Science in Industrial Engineering

In

The Department of Construction Management and Industrial Engineering

By
Sean Najm Ghazavi
B.S. Industrial and Systems Engineering,
Virginia Polytechnic Institute and State University, 2004
December 2007

Acknowledgments

I would like to express my deepest gratitude to my Advisor, Dr. T. Warren Liao. His constant guidance, encouragement and trust in my abilities helped me carve my way to this thesis research. His professional and friendly approach taught me many invaluable lessons. I would also like to thank Dr. Bin Li and Dr. Jing Wang for their comments and inputs in my work.

I would like to thank my lovely wife, Shayesteh, - who this work is dedicated to - for being there for me and supporting me through the process of finishing this work. Without her patience and encouragement this work would not have been completed. I would also like to thank my parents, brother and sister for their support.

A big thanks to all my friends who helped and supported me throughout. I would especially like to thank Hessam Babaei and Dr. Amin Shah-Hosseini for their valuable advice on scientific and academic matters. Additionally, I would like to convey sincere appreciation to Dr. Fereydoun Aghazadeh for his guidance throughout my Masters degree.

Table of Contents

Acknowledgments.....	ii
Abstract.....	iv
1. Introduction	1
2. Research Objective	4
3. Literature Review	5
3.1 Development of Automated Weld Inspection Systems.....	5
3.2 Feature Selection and Weld Inspection	8
4. Methodology.....	9
4.1 Background	9
4.2 Fuzzy Clustering	12
4.3 Fuzzy Modeling and Reasoning.....	17
4.4 Feature Selection	20
4.4.1 Mutual Correlation	22
4.4.2 Mutual Information	25
4.4.3 Stepwise Selection Method	27
4.4.4 Welch t -Statistics and Fisher Correlation Score.....	28
4.4.5 Relief Algorithm	30
5. Results and Discussion	31
5.1 Data Description	31
5.2 Test Results	31
6. Conclusion.....	53
References	55
Appendix	
A. Results of Breast Cancer Dataset.....	61
B. Results on PIMA Diabetes Dataset	73
Vita	85

Abstract

This thesis develops a computer-aided weld inspection methodology based on fuzzy modeling with selected features. The proposed methodology employs several filter feature selection methods for selecting input variables and then builds fuzzy models with the selected features. Our fuzzy modeling method is based on a fuzzy c-means (FCM) variant for the generation of fuzzy terms sets. The implemented FCM variant differs from the original FCM method in two aspects: (1) the two end terms take the maximum and minimum domain values as their centers, and (2) all fuzzy terms are forced to be convex. The optimal number of terms and the optimal shape of the membership function associated with each term are determined based on the mean squared error criterion. The fuzzy model serves as the rule base of a fuzzy reasoning based expert system implemented. In this implementation, first the fuzzy rules are extracted from feature data one feature at a time based on the FCM variant. The total number of fuzzy rules is the product of the fuzzy terms for each feature. The performances of these fuzzy sets are then tested with unseen data in terms of accuracy rates and computational time.

To evaluate the goodness of each selected feature subset, the selected combination is used as an input for the proposed fuzzy model. The accuracy of each selected feature subset along with the average error of the selected filter technique is reported. For comparison, the results of all possible combinations of the specified set of feature subsets are also obtained.

1. Introduction

In the fast paced world we live in today, there is not much room for errors. Any small mistake or flaw can cause catastrophic accidents in the engineering field. To prevent such incidents much research has been conducted in different areas; for instance, in design and development areas, in quality assurance, and many more. One way to prevent releasing any defective product into the system is to test it before sending it to the public. There are different ways of testing products. The two most prominent approaches are destructive testing and non-destructive testing. In destructive testing, tests are carried out till the specimen's failure. On the other hand, nondestructive, is testing that does not destroy the test object. The focus of this work is on the later as some welded structures should be tested nondestructively, particularly for critical applications where weld failure can be disastrous, such as in pressure vessels and power plants [1]. Some typical weld defects that need to be found and repaired are lack of fusion of the weld to the metal and porous bubbles inside the weld. Both of these could cause a structure to break or a pipeline to rupture.

Different methods for nondestructive testing of welded joints include visual examination, use of magnetic particles and/or liquid penetrant, ultrasonic, and radiography. Visual inspection is the oldest technique and extensively used to evaluate the condition of a weld. It can easily be preformed, is inexpensive and usually doesn't require special equipment. On the other hand, it requires good vision, good lighting and the knowledge of what to look for. Apart from the visual inspection, liquid penetration inspection (PT) is the oldest and most widely used nondestructive testing method. It is used to reveal surface breaking flaws by capillary action principles. Test objects are coated with visible or fluorescent dye solution. Excess dye is then removed from the surface, and a developer is applied. The developer acts as a blotter, drawing trapped penetrant out of imperfections open to the surface. With

visible dyes, vivid color contrasts between the penetrant and developer make cracks easy to see. It can detect pinholes and surface cracks which are not visible with naked eyes. PT can be used on any material and is most often used on materials clad in stainless steel, and stainless welded items which cannot be inspected by other methods [2]. Radiographic inspection has been one of the most widely used nondestructive techniques for detecting internal weld flaws in industries. Radiographs are exposed photographic images produced by permitting the X-ray or γ -ray source to penetrate the welds being inspected. Then these images are viewed by certified inspector in order to decide about the acceptability of the welds. Before any decision is made the inspector has to make sure that the image has the satisfactory quality as this is a key factor in weld flaw detection. Also, the inspector must have a solid understanding of the following in order to be successful with the interpretation of welding quality:

- Principles of radiographic examinations.
- Welding processes, their associated flaws, and their images as they appear on the radiograph.
- The acceptance criteria as specified in the codes and standards.

Training and certifying the film interpreter is time-consuming and expensive. Furthermore, these interpretations of weld quality based on radiographic images are inconsistent, subjective, sometimes biased, and very labor intensive. Therefore, it is of interest to design and develop a more robust approach by developing a computer-aided system to assist the human interpreter [3]. A considerable amount of research exists in the weld inspection area and will be explained in more detail in chapter 2. In developing a computer-aided weld inspection system several crucial steps have to be carried out. First, it is needed to develop some algorithms to extract the weld from the images. Then the next step is to detect the flaws. Finally, different types of welding flaws can be classified. The first step, extracting the weld from the image, has been done by applying the methodology carried out by Liao in his work of

“Fuzzy reasoning base inspection of radiographic welds: weld recognition”. In that particular study, a fuzzy expert system for segmentation of the welds was presented. It involves implementing the fuzzy c-means variant algorithm and then extracting fuzzy rules as the two major components of the fuzzy expert system for weld recognition [4].

The next step is to detect the flaws in the welds. Since a fuzzy expert system has already been developed, it is sensible to use it once more for this step as well. However, detecting flaws in welds require more features in order to detect all types of defects. Having more features in the model can be translated into a bigger model. Therefore, a feature selection method has to be implemented. Feature selection is the technique, commonly used in machine learning, for selecting a subset of relevant features for building robust learning models. By removing most irrelevant and redundant features from the data, feature selection might also help improve the performance of learning models. There are two main approaches for feature selection: filter method and wrapper method. The main interest of this paper lies in the filter technique, as it has not been applied for weld inspection to this day. Each technique, filter and wrapper, has its own advantages and disadvantage which will be discussed later in chapter 4. This research will not study the classification of different types of welding flaws. Theoretically, the methodologies proposed in this research can be easily adapted for that by extending the 2-class problem to the multi-class problem.

The rest of this thesis is summarized as follows: In chapter 2 the research objective is stated. Chapter 3 reviews related works. In chapter 4 the proposed methodology and fuzzy modeling are described. Chapter 4 also discusses different feature selections techniques to be employed in this research. Results and discussions are presented in chapter 5. And finally concluding remarks are presented.

2. Research Objective

There has been very limited research on applying feature selection techniques to automatic weld inspection systems. This research intends to make a contribution to this area by combining feature selection techniques and fuzzy expert systems to develop a new model which pre-process the data to obtain a fairly good subset of features and use them as an input to the fuzzy expert model. Feature selection ought to reduce the size of the model so as to make the model more comprehensible and to obtain the results in a shorter period of time, and increase the detection accuracy if possible.

This computer aided welding flaw detection system using selected features is realized by:

1. **Feature selection**
2. **Model learning and pattern classification**
3. **Automatic decision making**

Feature selection finds the optimized subset of features which describes the characteristics of welding defects. There are many features extracted from the images, however, only good feature selection could lead to higher efficiency and accuracy. Pattern classification analyzes historical data and makes prediction on the future data. Various pattern classification algorithms exist and each has its own advantages and disadvantages. In this research fuzzy reasoning based expert system is adopted. To extract the fuzzy rules, the fuzzy c-means variant based fuzzy modeling approach as proposed by [4] is applied. Automatic decision making is the final step as after identifying the weld, it has to be decided if it is a correct weld or a defect and if it is a defect how to categorize it based on the previous knowledge. The higher the accuracy, the better a model is learned.

3. Literature Review

This section has been divided into two parts. The first part outlines the works related to the development of automatic weld inspection system, and the second part reviews works related to the feature selection and weld inspection area.

3.1 Development of Automated Weld Inspection Systems

Several studies have been carried out to develop automated weld inspection systems. Here the emphasis is on the development of automated radiographic weld inspection. Several efforts have been devoted in order to achieve this. In general three main functions were investigated: Segmenting welds from the background, identifying the flawed segments in the weld, and classifying different types of welding flaws. The previously mentioned developments rely mostly on image processing, pattern recognition, feature extraction and pattern classification techniques. The later technique includes artificial neural network, rule-based reasoning, fuzzy c-means and fuzzy k-nearest neighbors.

Daum *et al.* [5] in 1987 developed an algorithm to segment the defects and mark them in the image. The detection was done by a background subtraction algorithm. Even though it was proven that the algorithm works independently of flaw types, it had difficulties detecting small defect regions (4 to 6 pixels). Gayer *et al.* [6] described an automatic welding defect recognition using radiography by a two-step process. This method was designed to mimic the human inspector by first going over the radiograph with a coarse resolution and then by fine focusing on the defective areas. Murakami [7] proposed a local arithmetic operation to a limited region and was followed by thresholding methods. He classified defect types with an expert system using information such as shape, position and intensity level of the defect pattern. The system could easily detect blowholes, but not cracks. Kato *et al.* [8] used

the expert system approach for identifying different types of welding flaw. They used 10 features to classify crack, lack of fusion, lack of penetration, porosity and inclusion defects, and extracted six features from each welding defect. The choice of the relevant features to be used in the system was based on interviews made with expert radiograph inspectors. It was concluded that it was very subjective to choose the features as each inspector adopted specific features of shape or geometry of the defect for categorizing. Hyatt *et al.* [9] demonstrated a multi-scale method designed to remove the overall background structure while reserving the defect details. Liao and Ni [1] developed a methodology to extract the welds from digitized radiographic images. This method is based on the resemblance of the distribution of pixels' intensities in the weld area to the Gaussian distribution. Each object was identified by three features: width, peak intensity and a term called mean square error (MSE). MSE is the similarity measure between the intensity profiles of the object to its Gaussian curve, calculated in terms of mean square error. This method has been proven to be effective; however, it is only capable of dealing with linear welds. Liao and Tang [10] presented a multilayered perceptron (MLP) based procedure for extracting welds from digitized radiographic images. The procedure consists of three major components: feature extraction, MLP-based object classification, and post-processing. This method can be applied to both linear and curved welds. The procedure is intended to extract welds before applying flaw detection algorithms. In this work, neural network was used as the pattern classifier. Liao and Li [11] presented another automatic radiographic nondestructive testing (NDT) system for weld inspection. The flaw detection methodology was developed based on the fitted line profiles of a weld image. In other words, it was based on the distortion in the overall line profile of the weld due to welding flaws. The process consisted of four modules: preprocessing, curve fitting, profile-anomaly detection, and post-processing. The results show that the system has a high detection rate and an acceptable false alarm rate. Aoki and Suga [12] used a three layer artificial neural network to identify automatically generated defect by image processing techniques based on ten discriminative features.

The algorithm achieved a successful rate of over 90% by using these discrimination features. Liao *et al.* [13] presented a weld flaw detection methodology based on fuzzy classification method. Two fuzzy classification methods, namely fuzzy k-nearest neighbor and fuzzy c-means, were applied and their performances were compared. It was shown that fuzzy k-nearest neighbor outperforms fuzzy c-means. Wang and Liao [14] applied the fuzzy KNN algorithm and MLP neural networks to classify six types of welding flaws using twelve features. Background subtraction and histogram thresholding were applied to segment the image. Their performances were tested and compared using the bootstrap method, and it was concluded that MLP outperformed fuzzy K-NN. Liao [15] applied fuzzy modeling to classify welding flaw types and develop an expert system. Two different methods, the WM method and Genetic algorithm, were used to generate fuzzy rules. WM is a well-known fuzzy modeling method proposed by Wang and Mandel [16]. Then the accuracy of these two methods were evaluated and compared with two other methods: fuzzy k-nearest neighbor and multi-layer perceptron neural network. The result indicated the fuzzy expert system outperformed the others in terms of accuracy. Kaftandjian *et al.* [17] presented an approach which combined Dempster-Shafer (DS) theory with fuzzy sets for improving automatic detection of weld defects. It consisted of modeling detection uncertainty in feature space by applying mass function weighted by membership degrees, and combining the features of objects using the DS combination rule. It was shown that by associating a confidence level to each detected object detecting the flaws can be more reliable and precise. Liao [18] developed a fuzzy reasoning based expert system for recognition of welds in radiographic images. First, each object in the image is identified and then described with a three-feature vector. Then the fuzzy rules are extracted from the feature data. These data are fed to the system one feature at a time to extract the rules based on a fuzzy c-mean variant algorithm. The numbers of fuzzy terms then are determined based on the mean square error criterion. The performance of the fuzzy expert system was shown to be better than that of MPL neural network. The shortcoming of this method is its speed because as the numbers of features are increased,

the time to calculate the results also increases. Wang and Wong [19] applied a segmentation methodology using the fuzzy c-means algorithm. First, quality of the image is enhanced by top-hat, bottom-hat filter and adaptive wavelet thresholding. Then the fuzzy c-means algorithm is applied to segment the radiographic image. The results demonstrate a good overall performance of segmenting the image from the background. Felisberto *et al.* [20] developed a new methodology for weld quality interpretation system. They extracted weld beads from radiograph images. Genetic algorithm was used to find suitable parameters values which matched the model image of a weld bead sample. The results showed that the proposed method was capable of checking the number of weld beads as well as the position, width, length, and angle of each weld bead with an accuracy of 94.4%.

3.2 Feature Selection and Weld Inspection

To the best of our knowledge, feature selection has not been widely applied to the automated weld inspection systems. Da Silva *et al.* [21] measured the relevance of a feature and then evaluated the performance of some selected features. Liao [22] applied feature selection to this problem by using a wrapper technique. In his approach two different ant colony optimization-based feature selection methods were applied. Each method follows a different search strategy and applies them to two different stages of computer-aided weld inspection. The performances of the ant colony based feature selections in term of classification error rate and CPU time were measured. In this research, we will apply other feature selection methods.

4. Methodology

Identifying an appropriate research approach and strategy is a crucial step in order to aid in the reasoning and flow of a research and to accomplish the research objective. In this section, the methodology used to conduct this research is explained. In order to do so a brief background of fuzzy modeling is first presented and discussed.

4.1 Background

The idea of fuzzy sets was first proposed by Zadeh [23]. The idea of fuzzy sets was consequently introduced to systems theory in 1973 [24], presenting a new class of systems called fuzzy systems. Zadeh [23] describes his proposed idea as “to provide a natural way of dealing with problems in which the source of imprecision is the absence of sharply defined criteria of class membership rather than presence of random variables”. Since his introduction of fuzzy systems, many successful applications of fuzzy sets and fuzzy systems have been observed in various areas. Fuzzy sets and systems’ attractiveness are due to the ability of capturing human thinking and understanding, besides their effectiveness in problem solving.

Early efforts of fuzzy modeling were mainly initiated by researchers in the fuzzy control area where fuzzy sets and fuzzy logic already had been successful. This method is mainly comprised of a set of If-Then rules. There are two major fuzzy modeling schemes:

- Linguistic models
- TSK models

Linguistic models are based on collection of If-Then rules with indistinct predictions and operated on Mamdani-like fuzzy reasoning [25], and TSK models are based on the Takagi-Sugeno-Kang (TSK) approach of reasoning [26, 27]. Among the two methods, linguistic models are more appropriate for revealing human understandable knowledge from real world data.

Structural identification and parameter identification are two major aspects in the identification of a fuzzy model. The structure of a fuzzy model can be determined by the number of variables, the number and shape of fuzzy terms of each variable and the number of rules comprising the model. Guillaume [28] defined three necessary conditions for a set of fuzzy models to be interpretable as:

1. The fuzzy partition must be readable, in the sense that the fuzzy sets can be interpreted as linguistic labels.
2. The set of rules must be as small as possible.
3. The If-part of the rules should be derived from a subset of independent variables rather than the full set.

Liao [29] adds one more necessary condition to what Guillaume proposed to cover a larger scope, in which other than fuzzy If-Then rules, the fuzzy decision tree is also included. That condition is:

4. The fuzzy partitions for each variable should not be too many and commonly shared by all rules.

The fuzzy partitions and associated parameters should be data-driven and optimized, using some algorithms, to minimize performance degradation.

Every fuzzy modeling method must address two fundamental issues: how to represent the fuzzy concepts in the most appropriate granularity level, and how to use this model to derive conclusions. Most fuzzy modeling studies adopted an arbitrarily chosen form and assumed all variables have the same form except Shi *et al.* [30] in which they allowed the algorithm to select one of the six previously

defined membership functions for each variable used in rule extraction. The most advanced and sophisticated approach is to allow the algorithm to determine the optimal numbers of membership functions for each variable. It is necessary to note that even though most studies define the membership function for each variable separately, some researches define n-dimensional membership function for the whole input product space.

Every fuzzy modeling method has to check the performance of the generated fuzzy model. In order to derive the solution for test datum every method must employ a fuzzy inference. The basic components of a fuzzy inference method are:

- Pattern matching
- Aggregation of matching degrees
- Implication operation
- Aggregation of rule results
- Defuzzification

The method adopted in this research to generate fuzzy If-Then rules is one of Mamdani-like form of linguistic models. Generally methods for the generation of Mamdani-like form of linguistic models can be divided into six categories as described by Liao [29] . They are

1. Grid partitioning
2. Fuzzy clustering
3. Genetic algorithm
4. Neural Networks
5. Hybrids
6. Others

For this research fuzzy clustering method is adopted to generate fuzzy If-Then rules. In the next section this method will be described in more details.

4.2 Fuzzy Clustering

Sugeno and Yasukawa [31] proposed the first fuzzy clustering-based method. They proposed to apply the fuzzy c-means (FCM) algorithm to only output data. The fuzzy c-means algorithm is a classical fuzzy clustering method belonging to the partitional clustering category. It is one of the unsupervised pattern recognition techniques. Unsupervised pattern classification does not have prior knowledge about any pattern. This issue can be addressed by clustering or self-organizing. The objective is to partition data in such a way that the data points within one cluster are as similar to each other as possible and as far away as it can be from the data point of other clusters. The number of clusters is either fixed or unknown. FCM is derived from the hard c-means algorithm. The hard c-means and its variants are all based on minimization of the sum of Euclidean distance between data and cluster centers, which indirectly minimizes the variance as follows:

$$\text{Min } J_1(\mathbf{U}, \mathbf{V}) = \sum_{i=1}^c \sum_{k=1}^n (u_{ik})^2 \|x_k - v_i\|^2 \quad (4.1)$$

Where x_k are data and $k = 1, \dots, n$, v_i are cluster centers and $i = 1, \dots, c$, $\mathbf{U} = [u_{ik}]$ denotes the matrix of hard c-partition, and $\mathbf{V} = \{v_i\}$ is the vector of all cluster centers. The partition constraints in c-means are:

1. $u_{ik} \in \{0,1\}, \forall i, k$
2. $\sum_{i=1}^c u_{ik} = 1, \forall k$
3. $0 < \sum_{k=1}^n u_{ik} < n, \forall i$

In other words, x_k belongs to one and only one cluster. Dunn [32] extended the fuzzy c-means algorithm from the hard c-means algorithm to allow for fuzzy partition with the objective function presented below:

$$\text{Min } J_2(\mathbf{U}, \mathbf{V}) = \sum_{i=1}^c \sum_{k=1}^n (\mu_{ik})^2 \|x_k - v_i\|^2 \quad (4.2)$$

Note that $\mathbf{U} = [\mu_{ik}]$ denotes the matrix of fuzzy c-partition. The fuzzy c-partition constraints are:

1. $\mu_{ik} \in [0,1], \forall i, k$
2. $\sum_{i=1}^c \mu_{ik} = 1, \forall k$
3. $0 < \sum_{k=1}^n \mu_{ik} < n, \forall i$

Therefore, each x_k could belong to more than one cluster to a fractional degree between 0 and 1. Bezdek [33] generalized $J_2(\mathbf{U}, \mathbf{V})$ to an infinite number of objective functions, i.e. $J_m(\mathbf{U}, \mathbf{V})$, where $1 \leq m \leq \infty$. This new objective function is still subject to the same constraints of fuzzy c-partition. Its equation is:

$$\text{Min } J_m(\mathbf{U}, \mathbf{V}) = \sum_{i=1}^c \sum_{k=1}^n (\mu_{ik})^m \|x_k - v_i\|^m \quad (4.3)$$

It is of importance to note that both hard c-mean and FCM algorithms try to minimize the variance of those data within each cluster. Now, by differentiating the objective function with respect to v_i (for fixed \mathbf{U}) and to μ_{ik} (for fixed \mathbf{V}) subject to above conditions, the following two conditions can be obtained:

$$v_i = \frac{\sum_{k=1}^n (\mu_{ik})^m x_k}{\sum_{k=1}^n (\mu_{ik})^m}, i = 1, \dots, c \quad (4.4)$$

$$\mu_{ik} = \frac{\left(\frac{1}{\|x_k - v_i\|^2} \right)^{1/(m-1)}}{\sum_{j=1}^c \left(\frac{1}{\|x_k - v_i\|^2} \right)^{1/(m-1)}}, \quad (4.5)$$

$$i = 1, \dots, c, k = 1, \dots, n$$

An iterative alternative optimization method is necessary to solve the fuzzy c-means model. The number of clusters, c , must be specified to run the procedure. This issue of predetermining the number of clusters is an inherent drawback of FCM and various validity measures have been proposed to determine this optimal number of clusters. In this research, as proposed by Liao *et al.* [4], the term with the lowest mean square error (MSE) defines the optimal number of clusters. Least MSE is also used to identify the most appropriate membership functions.

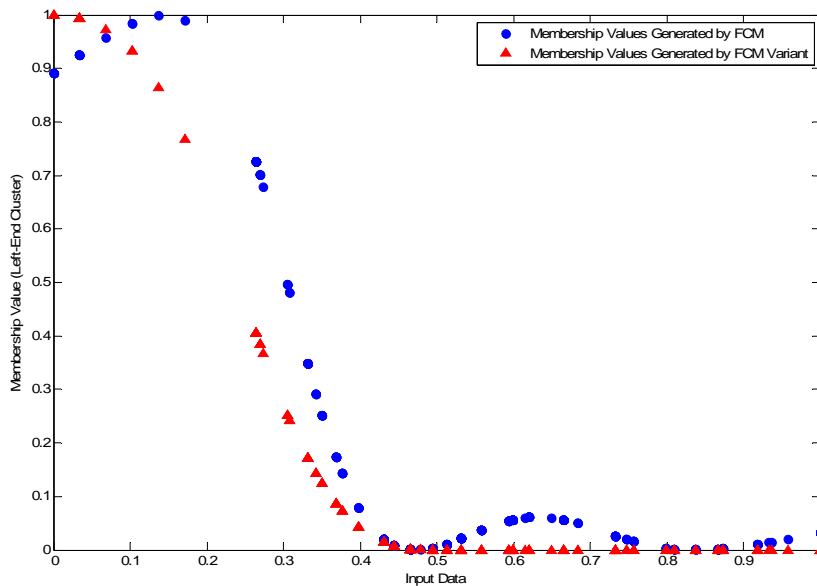
FCM has another drawback as pointed out by Medasani *et al.* [34] and confirmed by Liao *et al.* [13], Josien and Liao [35] and Liao *et al.* [3] which is sensitive to outliers. Therefore, clustering results generated from FCM cannot be trusted. It is expected, in general, a lower membership value to be associated with a domain value farther away from a term center, however, Liao *et al.* [4] showed otherwise. In other words, it was shown the generated membership function is not convex. In addition, it is also expected that for the left end term, a larger domain value to have a smaller membership value and the converse is true for the right end term. However, it was shown otherwise once again. As a result Liao *et al.* [4] proposed a modified FCM in which the above mentioned two problems are fixed. To achieve this original FCM was modified as follows:

- I. The lowest domain value should be set to be the center of the most left term (left end cluster) and the highest domain value to be the center of the most right term (right end cluster).
- II. Redistribute and normalize the concave part of a membership function to the other two more appropriate clusters.

Therefore, Liao *et al.* [4] proposed the following modified FCM algorithm which is adopted in this research as well:

- Step 1. Choose c ($2 \leq c \leq n$), m ($1 < m < \infty$), and ε (a small number as a threshold for stopping criterion). Set the counter $l = 0$ and initialize the membership matrix, $(U)^l$.
- Step 2. Set the highest domain value to be the center of the right-end cluster and lowest value to be the center of the left-end cluster.
- Step 3. Calculate the centers of the remaining clusters, $v_i^{(l)}$ using Equation 4.4.
- Step 4. Update the membership matrix $U^{(l+1)}$ by using Equation 4.5 if $x_k \neq v_i^{(l)}$. Otherwise, set $\mu_{ik} = 1(0)$ if $j = (\neq)i$.
- Step 5. Redistribute erroneous membership values to the two correct terms proportional to their current membership values.
- Step 6. Compute $\Delta = \|U^{(l+1)} - U^{(l)}\|$. If $\Delta > \varepsilon$, then increment l and go to step 3. If $\Delta \leq \varepsilon$, stop.

Each domain value is guaranteed to have a summation of membership value of one. A complete



membership to a certain fuzzy term simultaneously demonstrates a complete exclusion from all the remaining terms. Also, fuzzy terms so generated are convex and normal. Figure 1. shows the most left-end

Figure 1. Membership values generated by FCM and FCM variant on the same input variable

clusters generated by the original FCM and FCM variant out of three generated clusters (i.e., fuzzy terms). It is of importance to mention the modified FCM and FCM algorithms determines only the term centers. They do not explicitly define the shape of each fuzzy cluster. The fuzzy terms are asymmetrical as the term sets most likely are not uniformly spaced. To determine the shape, the data points of (d, m) pairs are fitted to some commonly used functions such as Triangular, Gaussian, s-shape, z-shape and π -shaped; where d represents domain value and m denotes membership value. The best shape of each fuzzy term is determined based on MSE value of each function. The best shape is the one with the lowest MSE value.

A triangular membership function, $TMF(x)$, is defined as:

$$TMF(x) = \begin{cases} \frac{x - \alpha}{\gamma - \alpha}, & \alpha \leq x \leq \gamma \\ \frac{x - \delta}{\gamma - \delta}, & \gamma \leq x \leq \delta \end{cases} \quad (4.6)$$

where $\alpha \leq \gamma \leq \delta$. Two adjacent fuzzy terms with triangular membership functions following the above assumptions should always have $\frac{1}{2}$ overlap at the midpoint between the two adjacent term centers.

A Gaussian membership function, $GMF(x)$, is defined as:

$$GMF(x) = e^{-\frac{1}{2}\left(x - \left(\frac{\mu}{\sigma}\right)\right)^2} \quad (4.7)$$

where μ and σ denote mean and standard deviation, respectively. Two adjacent fuzzy terms with Gaussian membership function with the same standard deviation should always have 0.223 overlap at the midpoint between the two term centers. Therefore, each domain value no longer has a total membership value of one. As a result, Gaussian membership function rarely is as suitable fit compared to other functions.

An s-shaped membership function, $SMF(x)$, is defined as:

$$SMF(x) = \begin{cases} 0, & x \leq \alpha \\ 2 \left(\frac{x - \alpha}{\gamma - \alpha} \right)^2, & \alpha \leq x \leq \beta \\ 1 - 2 \left(\frac{x - \gamma}{\gamma - \alpha} \right)^2, & \beta \leq x \leq \gamma \\ 1, & x \geq \gamma \end{cases} \quad (4.8)$$

where $\alpha \leq \beta \leq \gamma$. The complement of an SMF is a z-shaped function. An s-shaped and its associated z-shaped function which follow the above assumptions should always have $\frac{1}{2}$ overlap at the midpoint between the two term centers, i.e., $\beta = \frac{\alpha + \gamma}{2}$. A π -shaped function is a combination of s-shape and its associated z-shape function. In other words, it is comprised of an s-shaped function to its left side and a z-shaped function to its right side.

Now that the data is divided into different clusters, it is time to extract rules out of them. Next section will explain how to mine the If-Then rules from the clustered data.

4.3 Fuzzy Modeling and Reasoning

Different methods for fuzzy modeling in the literature were grouped into two broad trends by Delgado *et al.* [36]: The approximate approaches and the descriptive approaches. The first one, the approximate approach, tries to extract the fuzzy sets which characterize the fuzzy rules from the sample data without any intentions that the fuzzy sets have a linguistic interpretation. The descriptive approach, assumes that the linguistic values used in the antecedent and consequent of the fuzzy rules are taken from a set of fuzzy terms predefined in each domain of discourse. Let's consider a multiple input single output (MISO) system in which the A_i denotes i th input space and C stands for output are predefined set of terms. Assuming m input variables, the set of all possible rules may be presented by the Cartesian product:

$$R: A_1 \times \dots \times A_i \times \dots \times A_m \times C$$

The task is to select the best input space which better characterizes the system. There is one major problem associated with the descriptive approach which is as the number of input variables and/or the number of terms increases the combination of rules increases drastically. For instance, let's assume there is a data set which contains three inputs and one output. Then the fuzzy clustering method mentioned above is applied to each variable separately to divide each variable into appropriate number of terms. Let's assume again that the result of clustering is as follows: A_1 is divided into 2 terms, A_2 into 4 terms, A_3 into 3 terms, and C into 2 terms then the total number of rules are going to be $2 \times 4 \times 3 \times 2 = 48$ rules. To address this problem Delgado *et al.* [36] proposed to work with the fuzzy sets directly defined in the product space of the input variables, X^m , and those defined in the output space, Y .

Taking the descriptive approach, the fuzzy rules reflected on in this study have the following form:

$$R_{ijk}^l: \text{If } x_1 \text{ is } A_i, x_2 \text{ is } A_j, \text{ and } x_3 \text{ is } A_k, \text{ then } y \text{ is } C_l [w_{ijk}^l]$$

where A_i, A_j, A_k , and C_l are fuzzy terms defined in the domain of discourse corresponding to x_1, x_2, x_3 , and y , respectively; and w_{ijk}^l is the weight of the rule that can be interpreted as a certainty value relative to the "goodness" of the rule. The above mentioned rules can also be generalized if there are more variables. However, as it was mentioned before, as the number of variables increases the number of rules increases drastically, therefore in this research the maximum number of input allowed to go into the model is three with the maximum number of clusters which can be assigned to each variable is set to five. This would limit the numbers of rules not to exceed $5 \times 5 \times 5 \times 2 = 250$ rules. In this research, as proposed by Liao [18], two methods to determine w_{ijk}^l are used, assuming n numbers of input-output:

a) Max-min composition:

$$w_{ijk}^l = \text{Max}_{t=1,n} \text{Min} \left(A_i(x_1^t), A_j(x_2^t), A_k(x_3^t), C_l(y^t) \right) \quad (4.9)$$

where n is the total number of training data.

b) Mean-min composition:

$$w_{ijk}^l = \sum_{t=1,n} \text{Min} \left(A_i(x_1^t), A_j(x_2^t), A_k(x_3^t), C_l(y^t) \right) / n \quad (4.10)$$

$A_i(x_1^t)$ denotes the membership value which is obtained after matching the value of the t^{th} cluster of the first input variable to the fuzzy term A_i . $A_j(x_2^t)$, $A_k(x_3^t)$, and $C_l(y^t)$ are defined in the same manner. Both of the aforementioned methods use the MIN operator, as their t-norm operator, to aggregate all the matching degrees. The difference between these two methods is Max-min composition uses only one membership value of the examples (clusters) with the highest aggregated matching degree, and Mean-min composition uses the average of aggregated matching degrees for examples (clusters). Other methods can be used to calculate the w_{ijk}^l by replacing the Min, Max and Mean operators with other aggregation operators.

In order to utilize the knowledge, in the form of fuzzy If-Then rules here, resided in the rule base to infer an output value for any new input, a fuzzy reasoning method is needed. A reasoning method usually has the following components: pattern matching, aggregation of antecedents, implication operation, and aggregation of rule outputs. In this research the fuzzy reasoning method proposed by Liao [18] is employed. In this simplified reasoning method weights of each rule obtained from Equation 4.9 and Equation 4.10 has been taken into account to compute the output \hat{y} .

$$\hat{y} = \frac{\sum_{i=1}^{t1} \sum_{j=1}^{t2} \sum_{k=1}^{t3} \sum_{l=1}^{t0} \text{Min}(A_i(x_1), A_j(x_2), A_k(x_3), w_{ijk}^l) * v_l}{\sum_{i=1}^{t1} \sum_{j=1}^{t2} \sum_{k=1}^{t3} \sum_{l=1}^{t0} \text{Min}(A_i(x_1), A_j(x_2), A_k(x_3), w_{ijk}^l)} \quad (4.11)$$

$x = (x_1, x_2, x_3)$ represents new input vector to be inferred and v_l is a singleton consequent value, which corresponds to the centroid of the consequent term, C_l , in rule R_{ijk}^l . $A_i(x_1)$, $A_j(x_2)$, and $A_k(x_3)$ denote the membership value which are obtained after matching the first, second, and third input

variable value to the fuzzy term A_i , A_j , and A_k , correspondingly. If the output variable, y , is not a singleton, then the v_l value can be calculated by the center of gravity defuzzification method:

$$v_l = \frac{\int_Y y C_l(y) dy}{\int_Y C_l(y) dy} \quad (4.12)$$

The min operator is used for aggregating and implicating of the membership values of all inputs. Weighted averaging operation, where the weight is the aggregated membership value, is used for aggregation of rule outputs. After computing \hat{y} , it has to be translated into some understandable output. In this case, \hat{y} has to be 0 or 1. Traditionally in the literature a cutoff ration of 0.5 has been considered, meaning if computed \hat{y} is less than 0.5 then it is 0 and 1 otherwise. In this thesis, an exhaustive search finds the best cutoff ratio by changing this value from 0.75 to 0.25 in increments of 0.01. Finally \hat{y} is computed with the best cutoff ration which yields the lowest error. Depending on the method used, Max-Min method or Mean-Min, to calculate w_{ijk}^l , different values for equation 4.11 are obtained. It is of importance to note that all of the above equations consider three input variables and one output variable. Nevertheless, they can be easily generalized to other situations. In the worst case scenario, $5 \times 5 \times 5 \times 5 \times 2 = 1250$ rules have to be generated if the numbers of input variables are increased from 3 to 4, the model size is increased by five times. As a result, there is a need to manage the size of the model. Therefore, feature selection is considered to reduce the number of input variables to maximum of three input variables.

4.4 Feature Selection

Feature selection is a preprocessing technique, commonly used in machine learning, for selecting a subset of relevant features for building robust learning models. It can reduce dimensionality, remove

irrelevant data, increase learning accuracy, and improve results comprehensibility. Feature selection can be divided into three different categories: the first type encompasses algorithms built into adaptive systems for data analysis (predictors), second type of algorithms are wrapped around predictors providing them subset of features and receiving their feedback, and the third type are algorithms independent of any predictors, filtering out features which are irrelevant or redundant and are not very useful in data analysis. The filter approach evaluates and selects feature subsets based on general characteristics of data, and some statistical analysis without employing any learning model. On the other hand, wrapper technique involves a learning model, and uses its performance as the evaluation criterion. Wrapper approach is known to be more accurate compared to filter technique and it is computationally more expensive as well. Hybrid approach, which is a combination of filter and wrapper technique, is designed to trade accuracy with computational speed by applying wrapper techniques to only those subsets preselected by filter technique.

Liao [22] implemented the ant colony optimization as a wrapper technique of feature selection together with four different learning models, and to the best of our knowledge, no implementation of filter approach exists in the area of welding flaw detection. Therefore, it is proposed here to apply the filter approach of feature selection. There are different filter approaches in the literature which can be summarized as:

- Correlation based filtering
- Relevance indices based on distances between distributions
- Information theory
- Decision trees

For this research, use of correlation based filtering, mutual information, stepwise regression, relevance indices, and the Relief algorithm are proposed. In the following sections the details of each approach is described.

4.4.1 Mutual Correlation

Correlation is a well-known similarity measure indicating the strength and direction of a linear relationship between two random variables. If the variables are linearly dependent then their correlation coefficient is either +1 or -1. If the variables are uncorrelated or independent then their correlation coefficient is 0, but the converse is not true because the correlation coefficient detects only linear dependencies between two variables.

Three different mutual correlation indices were implemented in this research. One was proposed by Hiandl *et al.* [37], let's assume an n d -dimensional feature vectors

$$X_i = [^i x_1, \dots, ^i x_d] \quad i = 1, \dots, n$$

from c possible classes. Therefore, the mutual correlation for a feature pair of x_i and x_j is specified as:

$$r_{x_i, x_j} = \frac{\sum_c {}^c x_i * {}^c x_j - n \bar{x}_i \bar{x}_j}{\sqrt{(\sum_c {}^c x_i^2 - n \bar{x}_i^2)(\sum_c {}^c x_j^2 - n \bar{x}_j^2)}} \quad (4.13)$$

It is necessary to note that if two features x_i and x_j are independent then their correlation coefficient is zero, i.e. $r_{x_i, x_j} = 0$. The next step is to calculate all mutual correlations for all feature pairs and then compute the average absolute mutual correlation of a feature of a total of δ features:

$$r_{j, \delta} = \frac{1}{\delta} \sum_{i=1, i \neq j}^{\delta} |r_{x_i, x_j}| \quad (4.14)$$

Then the feature with the largest average mutual correlation will be removed at each iteration step:

$$\alpha = \arg \max_j r_{j,\delta} \quad (4.15)$$

Thus after removing feature x_α from the feature set, it should be removed from the average absolute mutual correlation calculation, i.e.

$$r_{j,\delta-1} = \frac{\delta r_{j,\delta} - |r_{x_\alpha, x_j}|}{\delta - 1} \quad (4.16)$$

Therefore the proposed feature selection algorithm by Hiandl *et al.* [37] can be summarized as:

- Step 1. Initialize $\delta = d - 1$.
- Step 2. Discard feature x_α for α determined by (4.15).
- Step 3. Decrement $\delta = \delta - 1$, if $\delta < D$ return the resulting D dimensional feature set and stop.
Otherwise,
- Step 4. Recalculate the average correlations by using (4.16).
- Step 5. Go to step 2.

The second one computes the merit of a feature subset, s , which consists of k number of features as

$$Merit_s = \frac{k \bar{r}_{cf}}{\sqrt{k + k(k-1) \bar{r}_{ff}}}, \quad (4.17)$$

where \bar{r}_{cf} and \bar{r}_{ff} denote the average feature-class correlation and feature-feature correlation, respectively [38]. The third one implements the mutual correlation feature selection method as proposed by Tsai and Chiu [39]. In their work, the mutual correlation between each input variable and out variable is obtained applying the following equation:

$$\rho(X, Y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \times \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (4.18)$$

Where \bar{x} and \bar{y} are the means of X and Y respectively. Instead of concerning positive or negative correlation, both situations are referred to as “mutual correlated”. Therefore, a mutual correlation measure between variables X and Y can be defined as $\rho(X, Y)^2$ where $0 \leq \rho(X, Y)^2 \leq 1$. If $\rho(X, Y)^2$ is close to 1, X and Y are strongly correlated; if $\rho(X, Y)^2$ is close to zero, X and Y are independent.

In this proposed feature selection method, a parameter p_i in the set P_r will be placed into the set P_r' , if $\rho(p_i, q)^2$ value between p_i and corresponding production quality q is high, as well as all $\rho(p_i, p_j)^2$ values between p_i and each principal parameter p_j in P_r' are low. Therefore, the selection priority for each parameter p_i in the parameter set P_r is defined as:

$$Prio(p_i) = \rho(p_i, q)^2 - \sum_{p_j \in P_r'} [\rho(p_i, q)^2 \times \rho(p_j, p_i)^2] \quad (4.19)$$

If $Prio(p_i) > Prio(p_k)$ for $\forall k \neq i$, the parameter p_i will be taken out from P_r and be placed into P_r' .

Their proposed algorithm can be summarized as:

Input : the parameter set P_r which contains all Z parameters, the production quality q

Output : the parameter set P_r' which contains T parameters taken out from P_r

Method :

- (1) Initialization : $P_r = \{p_i | i=1, \dots, Z\}$ and $P_r' = \emptyset$
- (2) Do until (P_r' includes T principal parameters) {
- (3) $\forall p_i \in P_r$, calculate the mutual correlation $\rho(p_i, q)^2$
- (4) $\forall p_i \in P_r$ and $\forall p_j \in P_r'$, calculate the mutual correlation $\rho(p_i, p_j)^2$
- (5) $\forall p_i \in P_r$, calculate selection priority $Prio(p_i)$ using Equation (5)
- (6) Sort P_r in decreasing value of $Prio(p_i)$, $\forall p_i \in P_r$
- (7) Select p_1 with maximal $Prio(p_i)$ in P_r
- (8) Set $P_r \leftarrow P_r - \{p_1\}$ and $P_r' \leftarrow P_r' + \{p_1\}$
- (9) }
- (10) Return P_r' // $P_r' = \{p_j | j=1, \dots, T\}$

4.4.2 Mutual Information

Mutual information, in probability theory and information theory, measures the amount of information which can be obtained about one random variable by observing another. Intuitively, mutual information measures the information that two discrete random variables X and Y share, therefore, it is the same as uncertainty contained in Y (or X) alone, namely the entropy of Y (or X). The entropy of variable X is defined as:

$$H(X) = - \sum_i P(x_i) \log_2(P(x_i)) \quad (4.20)$$

where $P(x_i)$ is the prior probability for all values of X . The joint entropy of two discrete random variables X and Y is merely the entropy of their pairing: (X,Y) . This implies that if X and Y are independent, then their joint entropy is the sum of their individual entropies, and it is defined as:

$$H(X,Y) = - \sum_i P(x_i, y_i) \log_2(P(x_i, y_i)) \quad (4.21)$$

The entropy of X after observing values of another variable Y is called conditional entropy. The conditional entropy of X given random variable Y (or equivocation of X about Y) is the average conditional entropy over Y :

$$H(X|Y) = - \sum_j P(y_j) \sum_i P(x_i|y_j) \log_2(P(x_i|y_j)) = \sum_{x,y} P(x,y) \log_2 \frac{P(y)}{P(x,y)} \quad (4.22)$$

where $P(x_i|y_j)$ is the posterior probabilities of X given values of Y . The mutual information of X relative to Y is given by:

$$I(X;Y) = \sum_{x,y} P(x,y) \log_2 \frac{P(y)}{P(x,y)} \quad (4.23)$$

Additional information about X provided by Y is directly related to the amount by which the entropy of X changes, which is called information gain [40]. Information gain is defined as:

$$IG(X|Y) = I(X; Y) = H(X) - H(X|Y) \quad (4.24)$$

It can be concluded from equation (4.21) that feature Y is more correlated to feature X than feature Z if $IG(X|Y) > IG(X|Z)$.

In this research the approach proposed by Peng *et al.* [41] is applied. They proposed a feature selection method based on mutual information. The criteria encompassed by Peng *et al.* [41] are maximum relevance and minimum redundancy of features based on mutual information. It should be noted mutual information works on discrete values. Therefore, for real data a discretization process must first be carried out before this mutual information based feature selection method can be used. In this research, two most basic discretization methods, i.e., equal width and equal frequency, are employed.

Let S denote the subset of features we are seeking. The minimum redundancy condition is:

$$\min W_I, W_I = \frac{1}{|S|^2} \sum_{i,j \in S} I(i, j) \quad (4.25)$$

where $|S| (= m)$ is the number of features in S . The maximum relevance condition is to maximize the total relevance of all features in S :

$$\max V_I, V_I = \frac{1}{|S|^2} \sum_{i \in S} I(h, i) \quad (4.26)$$

where $I(h, i)$ is the mutual information between targeted classes $h = \{h_1, h_2, \dots, h_k\}$, the classification variable, and the feature expression.

In another work Ding and Peng [42] elaborated more on the idea of maximum relevance, minimum redundancy, and developed two criteria:

- MID: Mutual Information Difference criterion
- MIQ: Mutual Information Quotient criterion

MID is the maximum of difference between relevancy measures and redundancy measures over all of the input variables, and MIQ is the maximum of relevancy measures divided by the redundancy measures over all of the input variables. In this thesis both methods have been incorporated.

4.4.3 Stepwise Selection Method

“Stepwise regression procedures are selection techniques that sequentially add or delete single predictor variables to the prediction equation” [43]. Before obtaining a final equation, a series of steps are taken and since each step directly leads to the next, these methods require a much smaller number of equations compared to the 2^n possible regressions approach. Major limitation of stepwise selection methods is that the end result is only a single subset that is not necessarily best for a given size of variables and few alternative subsets are suggested. There are three main approaches of stepwise selection methods:

1. Forward selection method: starting with no variable, and then adds predictor variables one at a time to the prediction equation if they are “statistically significant”. This process continues until either all predictors are included or some selection criterion is satisfied.
2. Backward elimination method: begins with all predictor variables included in the prediction equation, testing them one by one for statistical significance, and removing any which is not significant (produces smallest sum-of-square-error (SSE)). This procedure is continued until all predictors are excluded or some selection criterion is satisfied.
3. Stepwise procedure: is a combination of forward selection method and backward elimination technique. It is essentially a forward selection procedure but at each step of the process the

predictor variables in the chosen subset are re-examined for possible exclusion as it was done in the backward elimination method.

In this research the stepwise procedure is applied and subsets of variables are obtained. Since this subset of features still contains too many variables, only the three/four features which have highest coefficients and smallest confidence intervals are chosen as the variables going into the model.

4.4.4 Welch t -Statistics and Fisher Correlation Score

In this research, the Welch's t -statistics implemented by Loo *et al.* [44] for gene selection is also applied.

Welch's t -test is an adaptation of Student's t -test for two samples with unequal variances. It is defined as:

$$WTS(G_i) = \frac{|\mu_i^+ - \mu_i^-|}{\sqrt{\frac{(\sigma_i^+)^2}{n^+} + \frac{(\sigma_i^-)^2}{n^-}}} \quad (4.27)$$

Here n^+ and n^- are the numbers of expression levels, μ_i^+ and μ_i^- are the means of expression levels, and $(\sigma_i^+)^2$ and $(\sigma_i^-)^2$ are the variances of expression levels in the positive and negative classes, respectively. The vector G_i represents the gene expression levels of the i^{th} gene in all n samples. We use $y_j \in \{-1, +1\}$ to label the class of the j^{th} sample. After obtaining the test statistic for each gene, the statistical significance of each test statistic is assessed in order to identify differentially expressed genes. The assessment requires the distribution of the test statistic under null hypothesis. Same principle is applied to our weld test as well.

Fisher correlation score is a variant of Welch's t -statistics, which is defined as:

$$FCS(G_i) = \frac{|\mu_i^+ - \mu_i^-|}{\sigma_i^+ + \sigma_i^-} \quad (4.28)$$

WTS and its variants can be considered signal-to-noise ratio measurements of the expression levels of the input data, which in Loo *et al.*'s [44] work are genes.

When high-sensitivity screening is required, Loo *et al.* [44] showed WTS falls behind Average Distance Score (ADS) and Mean Distance Score (MDS). It is proposed in their work to replace the serial noise estimator $(\sigma_i^+ + \sigma_i^-)$ with a parallel noise estimator $\left(\frac{\sigma_i^+ \sigma_i^-}{\sigma_i^+ + \sigma_i^-}\right)$. MDS and ADS defined as follow:

$$ADS(G_i) = \frac{d_i^+ - d_i^-}{\left(\frac{\sigma_i^+ \sigma_i^-}{\sigma_i^+ + \sigma_i^-}\right)} \quad (4.29)$$

$$MDS(G_i) = \frac{|\mu_i^+ - \mu_i^-|}{\left(\frac{\sigma_i^+ \sigma_i^-}{\sigma_i^+ + \sigma_i^-}\right)} \quad (4.30)$$

where d_i^+ and d_i^- are defined as average difference between expression levels from one class to the mean of expression levels from another class. They can be presented as:

$$d_i^+ = \frac{1}{n^+} \sum_{j=1}^n |x_{i,j} - \mu_i^-| \left(\frac{1 + y_i}{2}\right) \quad (4.31)$$

$$d_i^- = \frac{1}{n^-} \sum_{j=1}^n |x_{i,j} - \mu_i^+| \left(\frac{1 + y_i}{2}\right) \quad (4.32)$$

They suggest that ADS actually generalizes the independently consistent expression (ICE) discriminator that was proposed by Bijlani *et al.* [45], which is also used as one of the filter techniques. ICE can be represented as:

$$ICE = \frac{1}{\sigma_i^- n^+} \sum_{j=1}^n |x_{i,j} - \mu_i^-| \left(\frac{1 + y_i}{2}\right) + \frac{1}{\sigma_i^+ n^-} \sum_{j=1}^n |x_{i,j} - \mu_i^+| \left(\frac{1 + y_i}{2}\right) \quad (4.33)$$

4.4.5 Relief Algorithm

Relief searches the dataset for its two nearest neighbors: one from the same class (nearest hit) and one from another class (nearest miss). The key idea is to estimate the quality of features based on how well their values distinguish between instances that are near each other. For instance, consider a randomly selected instance X from a data set S with k features. Relief updates the quality estimation $W[A_i]$ for all the features A_i based on the values of difference function about X , H or nearest hit and M or nearest miss. This process is repeated m times where m is a user-predefined parameter.

In this research the algorithm proposed by Kira and Rebdell [46] is implemented. Their proposed algorithm is as follows:

Given m -desired number of sampled instances, and k -number of features,

1. Set all weights $W[A_i] := 0.0$;
2. For $j := 1$ to m do begin
3. Randomly select an instance X ;
4. Find nearest hit H and nearest miss M
5. For $i := 1$ to k do begin
6. $W[A_i] := W[A_i] - \frac{diff(A_i, X, H)}{m} + \frac{diff(A_i, X, M)}{m}$;
7. End
8. End

5. Results and Discussion

5.1 Data Description

The total number of records used in this research is 399 data records with each comprised of 25 numeric input features and one output variable in 2 different classes: good weld and flawed weld (labeled as 0 and 1). For more description of the welding flaw data please refer to [13]. In order to process these data, first the binary-class data was randomized, and then stratified sampling was used to generate 5-fold cross validation of data, and finally these data were applied into the above proposed fuzzy modeling methodology. To further evaluate the performance of the proposed methodology, the same procedure was also applied to two other data sets breast cancer dataset and Pima Indians diabetes dataset. The three datasets used in this study is summarized in Table 1 as follows:

Table 1. Summary of datasets used in testing

Dataset	Number of features	Number of records	Number of classes	References
Weld Flaw Identification	25	399	2	(Liao et al. 1999)
Wisconsin Breast Cancer	30	569	2	(Mangasarian and Wolberg, 1990)
PIMA Indian Diabetes	8	768	2	(Sigillito V., 1990)

The results are discussed in the following section.

5.2 Test Results

Since the number of records in hand is limited, only 399 data points, the data is tested using a 5-fold cross-validation. First, the order of the data is randomized and sorted by class and then four-fifth (80%) of the data of each class is selected as training data and the model is tested on the remaining one-fifth (20%) of each class, which the model has not been exposed to before. This method is called stratified

sampling. For each class, the randomized and sorted data is divided into the number of folds, e.g. five, equal sections. Then the first section, e.g. first fold, of the data is used as test data and the remaining 4 folds are used as training data. After the model is trained, the same approach is used but instead of the first section, the second section of data is used as the testing data, and remaining as the training data.

As explained before, the size of the model, number of features and number of training data are key factors affecting the accuracy of the model. The more features go into the model, the bigger the model is, and more training data is required. Roughly speaking, at least 10 data points are needed per rule generated in the model. Since obtaining data can be either too expensive, or too complicated, and some of the features are redundant and/or irrelevant, only three features at a time were used as input to the model. To find the best 3-feature subset, all 25 choosing 3 (= 2300) combinations of inputs for 5 different folds were tried. The same approach is also applied to the breast cancer dataset and Pima Indian diabetes dataset.

The exhaustive test results are sorted in ascending order based on the average errors of all five folds calculated according to equation 4.9 and equation 4.10. The results are given in Tables 2 and 3. Note that only the best 20 results are shown. The same experiments were also carried out by replacing the proposed fuzzy model with the ANFIS model [47] (Table 4), which is a well-known fuzzy model, and the fuzzy K-nearest neighbor algorithm [48] with three different K values (Tables 5-7), applying the same data.

Table 2. Average errors of five folds on weld data from EQ. 4.9

i	j	k	Average Error Bases on Equation (4.9)	Standard Deviation
2	8	13	13.4	3.5777
2	5	13	14	3.1623
1	8	13	14.2	3.8341
2	4	12	14.6	4.1593
2	11	13	15.2	3.4205
7	12	24	15.6	1.3416
2	7	19	15.6	4.7749
2	7	13	15.8	2.8636
2	4	8	15.8	4.7645
2	13	19	15.8	5.7619
7	19	24	16	1
2	13	20	16	2.1213
1	8	12	16	2.7386
2	13	21	16	4.6904
2	4	6	16	8.1548
2	12	20	16.2	2.5884
2	13	18	16.2	4.2661
2	13	17	16.4	3.2863
2	13	14	16.4	3.5777
2	3	12	16.4	6.1887

Table 4. Average errors of five folds on weld data from ANFIS

i	j	k	Average Error based on ANFIS	Standard Deviation
2	7	12	10.6	2.4083
2	13	25	10.6	3.5777
2	6	13	11	1.8708
2	7	13	11	2.7386
2	12	13	11	4.1231
2	13	21	11	4.8477
2	5	13	11.2	3.5637
2	12	14	11.2	3.6332
2	12	25	11.2	3.7014
2	8	13	11.2	4.5497
2	13	19	11.4	3.0496
2	8	12	11.4	3.3615
2	12	19	11.4	3.4351
2	6	12	11.8	1.9235
2	6	19	11.8	1.9235
2	12	22	11.8	2.2804
2	7	19	11.8	2.9496
2	9	12	11.8	3.4205
2	10	12	11.8	3.5637
2	13	23	11.8	3.5637

Table 3. Average errors of five folds on weld data from EQ. 4.10

i	j	k	Average Error Bases on Equation (4.10)	Standard Deviation
2	7	12	12.8	2.0494
2	5	6	12.8	3.2711
2	6	23	13.2	3.8987
2	6	22	13.4	3.2094
2	6	10	14.2	3.5637
7	12	15	14.4	4.3359
2	6	9	14.4	4.7223
2	6	12	14.6	3.4351
1	7	12	14.6	4.2778
2	4	6	14.8	2.7749
7	12	24	15	1.2247
2	6	7	15.2	1.6432
2	8	12	15.2	1.6432
1	8	12	15.2	4.3243
2	6	15	15.4	1.1402
2	9	12	15.6	2.9665
1	12	13	15.6	3.2094
1	6	19	15.6	4.6152
7	12	20	15.8	1.6432
2	16	25	15.8	1.9235

Table 5. Average errors of five folds on weld data from F5NN

i	j	k	Average Error based on Fuzzy KNN (N=5)	Standard Deviation
2	4	5	6.8	3.0332
2	7	13	7.2	2.1679
2	4	25	8.4	2.881
2	8	13	8.6	2.0736
2	5	8	8.6	2.7019
2	7	12	8.8	0.8367
2	5	13	8.8	2.6833
2	8	12	9.2	1.0954
2	4	8	9.2	2.2804
2	13	20	9.4	2.0736
2	5	7	9.6	1.8166
2	5	10	9.6	2.3022
2	13	16	9.6	3.7815
2	4	12	9.8	1.4832
2	6	13	9.8	1.6432
2	13	25	9.8	1.6432
2	7	25	9.8	2.1679
1	2	5	9.8	2.6833
2	4	7	10	1.8708
2	8	25	10.2	3.3466

Table 6. Average errors of five folds on weld data from F10NN

i	j	k	Average Error based on Fuzzy KNN (N=10)	Standard Deviation
2	4	5	7	2.7386
2	7	13	7.4	1.8166
2	7	12	8	1
2	8	13	8.4	1.8166
2	5	8	8.4	2.0736
2	4	25	8.8	2.3875
2	5	13	8.8	2.5884
2	13	16	9	3.7417
2	5	10	9.2	2.5884
2	6	13	9.4	1.1402
2	8	12	9.4	1.8166
2	4	8	9.4	3.1305
2	4	7	9.6	0.8944
2	5	6	9.6	1.1402
2	7	25	9.6	1.1402
2	13	25	9.6	1.6733
2	4	13	9.6	1.8166
2	5	7	9.6	1.8166
2	4	12	9.6	1.9494
2	13	20	9.6	2.3022

Table 7. Average errors of five folds on weld data from F15NN

i	j	k	Average Error based on Fuzzy KNN (N=15)	Standard Deviation
2	4	5	7	2.7386
2	7	13	7.2	1.6432
2	8	13	8	1.5811
2	7	12	8.2	0.8367
2	5	8	8.2	2.0494
2	5	13	8.6	2.7019
2	13	16	8.8	4.1473
2	13	20	9	2.6458
2	6	13	9.2	1.3038
2	4	25	9.2	1.9235
2	8	12	9.2	1.9235
2	5	10	9.2	2.5884
2	5	6	9.4	0.8944
2	7	25	9.4	1.5166
2	4	12	9.6	1.5166
2	4	8	9.6	2.0736
2	5	15	9.6	2.7019
2	4	7	9.8	1.0954
2	4	13	9.8	1.3038
2	5	7	9.8	1.9235

The same results are calculated for breast cancer dataset and Pima Indians diabetes dataset. These results can be found in the appendix.

The accuracy of the model based on 320 training data points and 79 testing data point on average for the best selected combination of feature subset is 83.04% using Equation 4.9 and 83.78% applying Equation 4.10 incorporating only three features out of a total of 25 features.

Applying the same procedures, accuracy of the model on the breast cancer dataset is 96.46% by using Equation 4.9, and 96.63% by Equation 4.10. PIMA diabetes dataset yields 75.55% and 77.78% accuracy using Equations 4.9 and 4.10 respectively.

These models were run on a Pentium Core 2 Duo with clock speed of 2.2 GHz and the CPU times are given in Tables 8-10.

Table 8. CPU time of different folds of Weld datasets

	CPU time for Weld Data (Fuzzy Model)	CPU time for Weld Data (ANFIS Model)	CPU time for Weld Data (Fuzzy KNN,N=5)	CPU time for Weld Data (Fuzzy KNN,N=10)	CPU time for Weld Data (Fuzzy KNN,N=15)
Fold 1	28.469 Minutes	27.663 Minutes	0.343 Minutes	0.350 Minutes	0.337 Minutes
Fold 2	26.675 Minutes	26.150 Minutes	0.350 Minutes	0.346 Minutes	0.346 Minutes
Fold 3	28.949 Minutes	37.256 Minutes	0.335 Minutes	0.353 Minutes	0.356 Minutes
Fold 4	28.753 Minutes	28.330 Minutes	0.351 Minutes	0.348 Minutes	0.347 Minutes
Fold 5	27.441 Minutes	37.611 Minutes	0.350 Minutes	0.344 Minutes	0.348 Minutes
Total Time	140.288 Minutes	157.011 Minutes	1.730 Minutes	1.741 Minutes	1.734 Minutes

Table 9. CPU time of different folds of Breast Cancer datasets

	CPU time for Breat Cancer Data (Fuzzy Model)	CPU time for Breat Cancer Data (ANFIS Model)	CPU time for Breat Cancer Data (Fuzzy KNN,N=5)	CPU time for Breat Cancer Data (Fuzzy KNN,N=10)	CPU time for Breat Cancer Data (Fuzzy KNN,N=15)
Fold 1	73.921 Minutes	113.939 Minutes	0.924 Minutes	0.932 Minutes	0.936 Minutes
Fold 2	79.106 Minutes	175.751 Minutes	0.948 Minutes	0.945 Minutes	0.956 Minutes
Fold 3	76.714 Minutes	187.102 Minutes	0.944 Minutes	0.943 Minutes	0.952 Minutes
Fold 4	75.863 Minutes	186.409 Minutes	0.954 Minutes	0.938 Minutes	0.945 Minutes
Fold 5	71.736 Minutes	178.200 Minutes	0.923 Minutes	0.948 Minutes	0.947 Minutes
Total Time	377.341 Minutes	841.401 Minutes	4.693 Minutes	4.706 Minutes	4.735 Minutes

After applying different proposed feature selection methods, their results are presented in the following tables. Since each feature selection method finds a different feature subset for a different fold, the comparison has to be done fold by fold, rather than all five folds together.

Each feature selection method has selected a three-feature subset as the input for each model. Stepwise feature selection only reduced the number of features from 25 features to 9 features. Since it is using hypothesis testing and making decisions based on the p-values, and our goal is to reduce the

feature subset to only three features, therefore, the three features with highest regression coefficients and lowest p-values are selected.

Table 10. CPU time of different folds of PIMA Diabetes datasets

	CPU time for PIMA Diabetes Data (Fuzzy Model)	CPU time for PIMA Diabetes Data (ANFIS Model)	CPU time for PIMA Diabetes Data (Fuzzy KNN,N=5)	CPU time for PIMA Diabetes Data (Fuzzy KNN,N=10)	CPU time for PIMA Diabetes Data (Fuzzy KNN,N=15)
Fold 1	1.330 Minutes	1.236 Minutes	0.025 Minutes	0.025 Minutes	0.025 Minutes
Fold 2	1.127 Minutes	1.469 Minutes	0.023 Minutes	0.025 Minutes	0.024 Minutes
Fold 3	1.263 Minutes	1.605 Minutes	0.024 Minutes	0.025 Minutes	0.025 Minutes
Fold 4	1.040 Minutes	1.130 Minutes	0.024 Minutes	0.025 Minutes	0.023 Minutes
Fold 5	1.263 Minutes	1.531 Minutes	0.024 Minutes	0.022 Minutes	0.025 Minutes
Total Time	6.025 Minutes	6.970 Minutes	0.120 Minutes	0.122 Minutes	0.122 Minutes

In the following tables, only the test results on the weld dataset are presented, and the results on the other two datasets are presented in the Appendix. Tables 11-13 give the performance of feature subsets selected by the three mutual correlation methods. Tables 14-29 give the performance of feature subsets selected by the mutual information method for two different discretization methods, four different numbers of discretized values, and two different maximum-relevance-minimum-redundancy criteria. Tables 31-35 give the performance of feature subsets selected by the Welch t statistic, Fisher correlection score, ICE, MDA, and ADS, respectively. Table 36 gives the performance of feature subsets selected by the Relief algorithm.

Table 11. Feature Selection with Mutual Correlation on Weld Data

FS: mutual correlation (Tsai and Chiu) Weld Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	7	6	16	18	14	14	14	15	79.75%	77.22%	82.28%	82.28%	82.28%	81.01%
Fold2	2	7	6	39	14	10	14	13	13	50.63%	82.28%	87.34%	82.28%	83.54%	83.54%
Fold3	2	7	6	10	15	10	12	12	11	87.34%	81.01%	87.34%	84.81%	84.81%	86.08%
Fold4	2	7	6	34	15	17	14	14	14	56.96%	81.01%	78.48%	82.28%	82.28%	82.28%
Fold5	2	7	6	14	21	16	9	12	11	82.28%	73.42%	79.75%	88.61%	84.81%	86.08%
Average Error										71.39%	78.99%	83.04%	84.05%	83.54%	83.80%

Table 12. Feature Selection with Mutual Correlation on Weld Data

FS: mutal correlation (Haindl et al) Weld Data (Liao)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	12	13	23	31	25	23	23	22	22	60.76%	68.35%	70.89%	70.89%	72.15%	72.15%
Fold2	12	13	23	28	27	18	23	23	22	64.56%	65.82%	77.22%	70.89%	70.89%	72.15%
Fold3	23	24	25	36	22	21	22	22	23	54.43%	72.15%	73.42%	72.15%	72.15%	70.89%
Fold4	6	8	24	27	21	24	19	19	20	65.82%	73.42%	69.62%	75.95%	75.95%	74.68%
Fold5	12	13	23	23	25	20	25	22	21	70.89%	68.35%	74.68%	68.35%	72.15%	73.42%
Average Error										63.29%	69.62%	73.16%	71.65%	72.66%	72.66%

Table 13. Feature Selection with Mutual Correlation on Weld Data

FS: mutal correlation (Park et al) Weld Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	1	2	6	28	19	12	20	19	19	64.56%	75.95%	84.81%	74.68%	75.95%	75.95%
Fold2	1	2	6	32	15	10	18	17	17	59.49%	81.01%	87.34%	77.22%	78.48%	78.48%
Fold3	1	2	6	12	13	11	20	20	20	84.81%	83.54%	86.08%	74.68%	74.68%	74.68%
Fold4	1	2	6	15	16	17	22	21	21	81.01%	79.75%	78.48%	72.15%	73.42%	73.42%
Fold5	1	2	6	17	16	15	19	18	19	78.48%	79.75%	81.01%	75.95%	77.22%	75.95%
Average Error										73.67%	80.00%	83.54%	74.94%	75.95%	75.70%

Table 14. Feature Selection with Mutual Information on Weld Data

FS: Mutual Information w/ Equal Width on Weld Data (w=5)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	6	15	37	17	13	20	20	20	53.16%	78.48%	83.54%	74.68%	74.68%	74.68%
Fold2	2	22	7	40	23	12	12	13	12	49.37%	70.89%	84.81%	84.81%	83.54%	84.81%
Fold3	2	6	15	29	15	13	19	18	18	63.29%	81.01%	83.54%	75.95%	77.22%	77.22%
Fold4	2	7	10	32	20	16	13	12	12	59.49%	74.68%	79.75%	83.54%	84.81%	84.81%
Fold5	2	7	10	30	19	11	11	10	11	62.03%	75.95%	86.08%	86.08%	87.34%	86.08%
Average Error										57.47%	76.20%	83.54%	81.01%	81.52%	81.52%

Table 15. Feature Selection with Mutual Information on Weld Data

FS: Mutual Information w/ Equal Width on Weld Data (w=10)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	10	22	44	24	18	28	28	28	44.30%	69.62%	77.22%	64.56%	64.56%	64.56%
Fold2	2	10	22	49	24	16	25	25	25	37.97%	69.62%	79.75%	68.35%	68.35%	68.35%
Fold3	2	10	22	47	21	12	30	30	30	40.51%	73.42%	84.81%	62.03%	62.03%	62.03%
Fold4	2	10	22	46	19	14	24	24	24	41.77%	75.95%	82.28%	69.62%	69.62%	69.62%
Fold5	2	10	22	39	19	15	29	29	29	50.63%	75.95%	81.01%	63.29%	63.29%	63.29%
Average Error										43.04%	72.91%	81.01%	65.57%	65.57%	65.57%

Table 16. Feature Selection with Mutual Information on Weld Data

FS: Mutual Information w/ Equal Width on Weld Data (w=15)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	10	22	44	24	18	28	28	28	44.30%	69.62%	77.22%	64.56%	64.56%	64.56%
Fold2	2	10	15	49	20	18	25	25	26	37.97%	74.68%	77.22%	68.35%	68.35%	67.09%
Fold3	2	10	22	47	21	12	30	30	30	40.51%	73.42%	84.81%	62.03%	62.03%	62.03%
Fold4	2	10	22	46	19	14	24	24	24	41.77%	75.95%	82.28%	69.62%	69.62%	69.62%
Fold5	2	10	22	39	19	15	29	29	29	50.63%	75.95%	81.01%	63.29%	63.29%	63.29%
Average Error										43.04%	73.92%	80.51%	65.57%	65.57%	65.32%

Table 17. Feature Selection with Mutual Information on Weld Data

FS: Mutual Information w/ Equal Width on Weld Data (w=20)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	10	15	48	19	19	24	23	23	39.24%	75.95%	75.95%	69.62%	70.89%	70.89%
Fold2	2	10	15	49	20	18	25	25	26	37.97%	74.68%	77.22%	68.35%	68.35%	67.09%
Fold3	2	10	15	32	17	16	24	24	24	59.49%	78.48%	79.75%	69.62%	69.62%	69.62%
Fold4	2	10	15	47	19	16	25	24	24	40.51%	75.95%	79.75%	68.35%	69.62%	69.62%
Fold5	2	10	15	45	17	15	27	27	27	43.04%	78.48%	81.01%	65.82%	65.82%	65.82%
Average Error										44.05%	76.71%	78.73%	68.35%	68.86%	68.61%

Table 18. Feature Selection with Mutual Information on Weld Data

FS: Mutual Information w/ Equal Frequency on Weld Data (f=5)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	10	12	2	25	18	15	20	20	18	68.35%	77.22%	81.01%	74.68%	74.68%	77.22%
Fold2	10	3	8	29	20	19	23	25	25	63.29%	74.68%	75.95%	70.89%	68.35%	68.35%
Fold3	10	12	2	21	18	9	16	16	17	73.42%	77.22%	88.61%	79.75%	79.75%	78.48%
Fold4	10	4	8	39	18	17	16	14	14	50.63%	77.22%	78.48%	79.75%	82.28%	82.28%
Fold5	10	2	12	14	14	8	11	11	11	82.28%	82.28%	89.87%	86.08%	86.08%	86.08%
Average Error										67.59%	77.72%	82.78%	78.23%	78.23%	78.48%

Table 19. Feature Selection with Mutual Information on Weld Data

FS: Mutual Information w/ Equal Frequency on Weld Data (f=10)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	10	11	2	27	23	19	24	24	25	65.82%	70.89%	75.95%	69.62%	69.62%	68.35%
Fold2	10	11	2	19	21	18	25	22	21	75.95%	73.42%	77.22%	68.35%	72.15%	73.42%
Fold3	10	19	2	21	21	14	16	17	16	73.42%	73.42%	82.28%	79.75%	78.48%	79.75%
Fold4	10	11	2	27	25	15	19	18	18	65.82%	68.35%	81.01%	75.95%	77.22%	77.22%
Fold5	10	3	12	20	13	20	18	18	19	74.68%	83.54%	74.68%	77.22%	77.22%	75.95%
Average Error										71.14%	73.92%	78.23%	74.18%	74.94%	74.94%

Table 20. Feature Selection with Mutual Information on Weld Data

FS: Mutual Information w/ Equal Frequency on Weld Data (f=15)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	10	11	4	34	34	30	21	21	21	56.96%	56.96%	62.03%	73.42%	73.42%	73.42%
Fold2	10	12	1	25	22	16	21	19	19	68.35%	72.15%	79.75%	73.42%	75.95%	75.95%
Fold3	10	11	2	33	16	16	19	18	17	58.23%	79.75%	79.75%	75.95%	77.22%	78.48%
Fold4	15	11	1	18	20	18	16	16	16	77.22%	74.68%	77.22%	79.75%	79.75%	79.75%
Fold5	10	4	12	22	16	18	14	13	13	72.15%	79.75%	77.22%	82.28%	83.54%	83.54%
Average Error										66.58%	72.66%	75.19%	76.96%	77.97%	78.23%

Table 21. Feature Selection with Mutual Information on Weld Data

FS: Mutual Information w/ Equal Frequency on Weld Data (f=20)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	10	1	12	28	23	16	21	23	23	64.56%	70.89%	79.75%	73.42%	70.89%	70.89%
Fold2	10	1	12	25	22	16	21	19	19	68.35%	72.15%	79.75%	73.42%	75.95%	75.95%
Fold3	10	12	4	26	20	15	15	17	17	67.09%	74.68%	81.01%	81.01%	78.48%	78.48%
Fold4	10	12	2	16	15	16	15	15	15	79.75%	81.01%	79.75%	81.01%	81.01%	81.01%
Fold5	10	25	12	23	26	19	14	14	15	70.89%	67.09%	75.95%	82.28%	82.28%	81.01%
Average Error										70.13%	73.16%	79.24%	78.23%	77.72%	77.47%

Table 22. Feature Selection with Mutual Information on Weld Data

FS: Mutual Information w/ Equal Width on Weld Data (w=5)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	6	7	16	18	14	14	14	15	79.75%	77.22%	82.28%	82.28%	82.28%	81.01%
Fold2	2	22	7	40	23	12	12	13	12	49.37%	70.89%	84.81%	84.81%	83.54%	84.81%
Fold3	2	6	15	29	15	13	19	18	18	63.29%	81.01%	83.54%	75.95%	77.22%	77.22%
Fold4	2	7	12	14	12	14	8	8	8	82.28%	84.81%	82.28%	89.87%	89.87%	89.87%
Fold5	2	7	12	13	11	9	10	9	9	83.54%	86.08%	88.61%	87.34%	88.61%	88.61%
Average Error										71.65%	80.00%	84.30%	84.05%	84.30%	84.30%

Table 23. Feature Selection with Mutual Information on Weld Data

FS: Mutual Information w/ Equal Width on Weld Data (w=10)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	10	12	25	18	15	20	20	18	68.35%	77.22%	81.01%	74.68%	74.68%	77.22%
Fold2	2	10	12	20	19	11	17	17	18	74.68%	75.95%	86.08%	78.48%	78.48%	77.22%
Fold3	2	10	7	20	19	13	12	11	11	74.68%	75.95%	83.54%	84.81%	86.08%	86.08%
Fold4	2	10	7	32	20	16	13	12	12	59.49%	74.68%	79.75%	83.54%	84.81%	84.81%
Fold5	2	10	7	30	19	11	11	10	11	62.03%	75.95%	86.08%	86.08%	87.34%	86.08%
Average Error										67.85%	75.95%	83.29%	81.52%	82.28%	82.28%

Table 24. Feature Selection with Mutual Information on Weld Data

FS: Mutual Information w/ Equal Width on Weld Data (w=15)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	22	12	25	21	14	16	16	16	68.35%	73.42%	82.28%	79.75%	79.75%	79.75%
Fold2	2	22	12	20	21	12	14	13	13	74.68%	73.42%	84.81%	82.28%	83.54%	83.54%
Fold3	2	22	7	10	15	8	14	14	13	87.34%	81.01%	89.87%	82.28%	82.28%	83.54%
Fold4	2	22	17	44	22	16	22	21	20	44.30%	72.15%	79.75%	72.15%	73.42%	74.68%
Fold5	2	12	7	13	11	9	10	9	9	83.54%	86.08%	88.61%	87.34%	88.61%	88.61%
Average Error										71.65%	77.22%	85.06%	80.76%	81.52%	82.03%

Table 25. Feature Selection with Mutual Information on Weld Data

FS: Mutual Information w/ Equal Width on Weld Data (w=20)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	10	12	25	18	15	20	20	18	68.35%	77.22%	81.01%	74.68%	74.68%	77.22%
Fold2	2	10	12	20	19	11	17	17	18	74.68%	75.95%	86.08%	78.48%	78.48%	77.22%
Fold3	2	10	12	21	18	9	16	16	17	73.42%	77.22%	88.61%	79.75%	79.75%	78.48%
Fold4	2	10	15	47	19	16	25	24	24	40.51%	75.95%	79.75%	68.35%	69.62%	69.62%
Fold5	2	10	15	45	17	15	27	27	27	43.04%	78.48%	81.01%	65.82%	65.82%	65.82%
Average Error										60.00%	76.96%	83.29%	73.42%	73.67%	73.67%

Table 26. Feature Selection with Mutual Information on Weld Data

FS: Mutual Information w/ Equal Frequency on Weld Data (f=5)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	10	12	2	25	18	15	20	20	18	68.35%	77.22%	81.01%	74.68%	74.68%	77.22%
Fold2	10	2	21	48	25	19	29	29	29	39.24%	68.35%	75.95%	63.29%	63.29%	63.29%
Fold3	10	12	2	21	18	9	16	16	17	73.42%	77.22%	88.61%	79.75%	79.75%	78.48%
Fold4	10	2	18	44	21	16	24	24	24	44.30%	73.42%	79.75%	69.62%	69.62%	69.62%
Fold5	10	2	18	44	21	16	28	27	26	44.30%	73.42%	79.75%	64.56%	65.82%	67.09%
Average Error										53.92%	73.92%	81.01%	70.38%	70.63%	71.14%

Table 27. Feature Selection with Mutual Information on Weld Data

FS: Mutual Information w/ Equal Frequency on Weld Data (f=10)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	10	2	15	48	19	19	24	23	23	39.24%	75.95%	75.95%	69.62%	70.89%	70.89%
Fold2	10	2	15	49	20	18	25	25	26	37.97%	74.68%	77.22%	68.35%	68.35%	67.09%
Fold3	10	2	15	32	17	16	24	24	24	59.49%	78.48%	79.75%	69.62%	69.62%	69.62%
Fold4	10	2	22	46	19	14	24	24	24	41.77%	75.95%	82.28%	69.62%	69.62%	69.62%
Fold5	10	2	18	44	21	16	28	27	26	44.30%	73.42%	79.75%	64.56%	65.82%	67.09%
Average Error										44.56%	75.70%	78.99%	68.35%	68.86%	68.86%

Table 28. Feature Selection with Mutual Information on Weld Data

FS: Mutual Information w/ Equal Frequency on Weld Data (f=15)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	10	2	23	38	22	19	28	28	28	51.90%	72.15%	75.95%	64.56%	64.56%	64.56%
Fold2	10	2	21	48	25	19	29	29	29	39.24%	68.35%	75.95%	63.29%	63.29%	63.29%
Fold3	10	2	23	44	21	13	28	28	28	44.30%	73.42%	83.54%	64.56%	64.56%	64.56%
Fold4	15	2	10	47	19	16	25	24	24	40.51%	75.95%	79.75%	68.35%	69.62%	69.62%
Fold5	10	2	23	39	17	15	30	30	30	50.63%	78.48%	81.01%	62.03%	62.03%	62.03%
Average Error										45.32%	73.67%	79.24%	64.56%	64.81%	64.81%

Table 29. Feature Selection with Mutual Information on Weld Data

FS: Mutual Information w/ Equal Frequency on Weld Data (f=20)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	10	2	18	46	25	20	27	27	27	41.77%	68.35%	74.68%	65.82%	65.82%	65.82%
Fold2	10	2	18	49	23	19	27	27	27	37.97%	70.89%	75.95%	65.82%	65.82%	65.82%
Fold3	10	18	6	49	17	17	19	19	19	37.97%	78.48%	78.48%	75.95%	75.95%	75.95%
Fold4	10	2	18	44	21	16	24	24	24	44.30%	73.42%	79.75%	69.62%	69.62%	69.62%
Fold5	10	2	15	45	17	15	27	27	27	43.04%	78.48%	81.01%	65.82%	65.82%	65.82%
Average Error										41.01%	73.92%	77.97%	68.61%	68.61%	68.61%

Table 30. Feature Selection with Stepwise on Weld Data

FS: Stepwise on Weld Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	6	10	37	18	14	28	28	27	53.16%	77.22%	82.28%	64.56%	64.56%	65.82%
Fold2	2	6	20	26	20	12	20	11	11	67.09%	74.68%	84.81%	74.68%	86.08%	86.08%
Fold3	2	6	15	29	15	13	19	18	18	63.29%	81.01%	83.54%	75.95%	77.22%	77.22%
Fold4	2	6	15	40	16	15	19	19	19	49.37%	79.75%	81.01%	75.95%	75.95%	75.95%
Fold5	2	6	20	18	14	14	12	13	11	77.22%	82.28%	82.28%	84.81%	83.54%	86.08%
Average Error										62.03%	78.99%	82.78%	75.19%	77.47%	78.23%

Table 31. Feature Selection with Welch t-Statistics on Weld Data

FS: Welch t-Statistics I on Weld Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	1	7	25	19	15	16	15	15	68.35%	75.95%	81.01%	79.75%	81.01%	81.01%
Fold2	2	1	7	30	22	14	16	15	16	62.03%	72.15%	82.28%	79.75%	81.01%	79.75%
Fold3	2	1	7	19	20	12	18	17	16	75.95%	74.68%	84.81%	77.22%	78.48%	79.75%
Fold4	2	1	7	22	17	19	16	16	17	72.15%	78.48%	75.95%	79.75%	79.75%	78.48%
Fold5	2	1	7	17	19	18	13	13	14	78.48%	75.95%	77.22%	83.54%	83.54%	82.28%
Average Error										71.39%	75.44%	80.25%	80.00%	80.76%	80.25%

Table 32. Feature Selection with Fisher Correlation Score on Weld Data

FS: Fisher Correlation Score on Weld Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	1	7	25	19	15	16	15	15	68.35%	75.95%	81.01%	79.75%	81.01%	81.01%
Fold2	2	1	7	30	22	14	16	15	16	62.03%	72.15%	82.28%	79.75%	81.01%	79.75%
Fold3	2	1	7	19	20	12	18	17	16	75.95%	74.68%	84.81%	77.22%	78.48%	79.75%
Fold4	2	1	7	22	17	19	16	16	17	72.15%	78.48%	75.95%	79.75%	79.75%	78.48%
Fold5	2	1	7	17	19	18	13	13	14	78.48%	75.95%	77.22%	83.54%	83.54%	82.28%
Average Error										71.39%	75.44%	80.25%	80.00%	80.76%	80.25%

Table 33. Feature Selection with Independently Consistent Expression (ICE) on Weld Data

FS: Independently consistent Expression on Weld Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	6	10	9	46	26	24	25	25	25	41.77%	67.09%	69.62%	68.35%	68.35%	68.35%
Fold2	6	10	15	49	24	19	23	23	23	37.97%	69.62%	75.95%	70.89%	70.89%	70.89%
Fold3	6	10	2	41	10	12	23	23	23	48.10%	87.34%	84.81%	70.89%	70.89%	70.89%
Fold4	6	15	2	40	16	15	19	19	19	49.37%	79.75%	81.01%	75.95%	75.95%	75.95%
Fold5	6	15	1	17	16	20	20	19	19	78.48%	79.75%	74.68%	74.68%	75.95%	75.95%
Average Error										51.14%	76.71%	77.22%	72.15%	72.41%	72.41%

Table 34. Feature Selection with Mean Difference Score (MDS) on Weld Data

FS: Mean Difference Score on Weld Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	1	7	25	19	15	16	15	15	68.35%	75.95%	81.01%	79.75%	81.01%	81.01%
Fold2	2	1	7	30	22	14	16	15	16	62.03%	72.15%	82.28%	79.75%	81.01%	79.75%
Fold3	2	1	7	19	20	12	18	17	16	75.95%	74.68%	84.81%	77.22%	78.48%	79.75%
Fold4	2	1	7	22	17	19	16	16	17	72.15%	78.48%	75.95%	79.75%	79.75%	78.48%
Fold5	2	1	7	17	19	18	13	13	14	78.48%	75.95%	77.22%	83.54%	83.54%	82.28%
Average Error										71.39%	75.44%	80.25%	80.00%	80.76%	80.25%

Table 35. Feature Selection with Average Difference Score (ADS) on Weld Data

FS: Average Difference Score on Weld Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	15	25	1	30	24	20	14	13	13	62.03%	69.62%	74.68%	82.28%	83.54%	83.54%
Fold2	9	25	3	20	22	17	16	18	18	74.68%	72.15%	78.48%	79.75%	77.22%	77.22%
Fold3	25	13	1	11	16	11	11	10	11	86.08%	79.75%	86.08%	86.08%	87.34%	86.08%
Fold4	18	14	15	40	28	23	28	28	28	49.37%	64.56%	70.89%	64.56%	64.56%	64.56%
Fold5	13	6	15	31	23	18	19	19	19	60.76%	70.89%	77.22%	75.95%	75.95%	75.95%
Average Error										66.58%	71.39%	77.47%	77.72%	77.72%	77.47%

Table 36. Feature Selection with Relief Algorithm on Weld Data

FS: Relief Algorithm on Weld Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	7	15	2	21	16	12	11	11	12	73.42%	79.75%	84.81%	86.08%	86.08%	84.81%
Fold2	1	15	17	46	28	22	21	20	20	41.77%	64.56%	72.15%	73.42%	74.68%	74.68%
Fold3	1	14	17	44	15	17	15	15	15	44.30%	81.01%	78.48%	81.01%	81.01%	81.01%
Fold4	1	7	2	22	17	19	16	16	17	72.15%	78.48%	75.95%	79.75%	79.75%	78.48%
Fold5	1	7	14	23	18	22	16	17	16	70.89%	77.22%	72.15%	79.75%	78.48%	79.75%
Average Error										60.51%	76.20%	76.71%	80.00%	80.00%	79.75%

The results of all six models are in the range of 70% to 85% accuracies for each fold, and 51% to 81% on average based on the various feature selection methods applied. Overall ANFIS is the most accurate of all three modeling methods yielding the highest average accuracy, in 19 out of 26 different feature selection results. The proposed fuzzy reasoning technique applying Equation 4.10 fairs closely with the fuzzy k-nearest neighbor algorithm with each outperforms the other depending on which feature selection method is applied. The fuzzy k-nearest method is the fastest method; however, it does not provide any explicit knowledge. Our fuzzy modeling method, on the other hand, has the advantage of generating human comprehensible model. ANFIS is generally slower and takes longer to compute the output. In our case, when number of records for weld dataset is small and only three features were used, calculation time of the ANFIS is comparable with our proposed model; however, when a bigger dataset, e.g. breast cancer dataset, is used ANFIS takes more than twice as long as our proposed model. Equation 4.9 is the less accurate because of using max operator as the aggregation operator. We believe mean operator works better than max operator due to the combination of results from different rules. Fuzzy k-nearest neighbor is the fastest method because the number of records is very limited, and fuzzy KNN algorithm searches the data almost exhaustively, meaning if $K=5$, therefore, there is $5 \times \text{number of records}$ searches it has to go through to calculate the output. Therefore, as the number of inputs is small fuzzy KNN is the fastest technique. This situation might change as the number of records increases.

Based on the average accuracy, the results of all feature selection-model combinations are ranked, as shown in Table 37 for the weld dataset. According to Table 37, the best result of 85.06% (in Table 24) was obtained by using the mutual information (discretization with equal width, $w=15$, and the miq criterion) for feature selection and the ANIFS fuzzy model. The second and third best results were obtained using the mutual information (discretization with equal width, $w=5$, and the miq criterion) for feature selection, and ANFIS and Fuzzy 10NN as their model respectively.

Table 37. Ranking of feature selection-model combination for the weld dataset

Rank	Average Accuracy	Modeling Method	Feature Selection
1	85.06%	ANFIS	FS: Mutual Information w/ Equal Width on Weld Data (w=15)(miq)
2	84.30%	ANFIS	FS: Mutual Information w/ Equal Width on Weld Data (w=5)(miq)
3	84.30%	Fuzzy 10NN	FS: Mutual Information w/ Equal Width on Weld Data (w=5)(miq)
4	84.30%	Fuzzy 15NN	FS: Mutual Information w/ Equal Width on Weld Data (w=5)(miq)
5	84.05%	Fuzzy 5NN	FS: Mutual correlation (Tsai and Chiu) Weld Data
6	84.05%	Fuzzy 5NN	FS: Mutual Information w/ Equal Width on Weld Data (w=5)(miq)
7	83.80%	Fuzzy 15NN	FS: Mutual correlation (Tsai and Chiu) Weld Data
8	83.54%	ANFIS	FS: Mutual correlation (Park et al) Weld Data
9	83.54%	ANFIS	FS: Mutual Information w/ Equal Width on Weld Data (w=5)(mid)
10	83.54%	Fuzzy 10NN	FS: Mutual correlation (Tsai and Chiu) Weld Data
11	83.29%	ANFIS	FS: Mutual Information w/ Equal Width on Weld Data (w=10)(miq)
12	83.29%	ANFIS	FS: Mutual Information w/ Equal Width on Weld Data (w=20)(miq)
13	83.04%	ANFIS	FS: Mutual correlation (Tsai and Chiu) Weld Data
14	82.78%	ANFIS	FS: Stepwise on Weld Data
15	82.78%	ANFIS	FS: Mutual Information w/ Equal Frequency on Weld Data (f=5)(mid)
16	82.28%	Fuzzy 10NN	FS: Mutual Information w/ Equal Width on Weld Data (w=10)(miq)
17	82.28%	Fuzzy 15NN	FS: Mutual Information w/ Equal Width on Weld Data (w=10)(miq)
18	82.03%	Fuzzy 15NN	FS: Mutual Information w/ Equal Width on Weld Data (w=15)(miq)
19	81.52%	Fuzzy 10NN	FS: Mutual Information w/ Equal Width on Weld Data (w=5)(mid)
20	81.52%	Fuzzy 15NN	FS: Mutual Information w/ Equal Width on Weld Data (w=5)(mid)

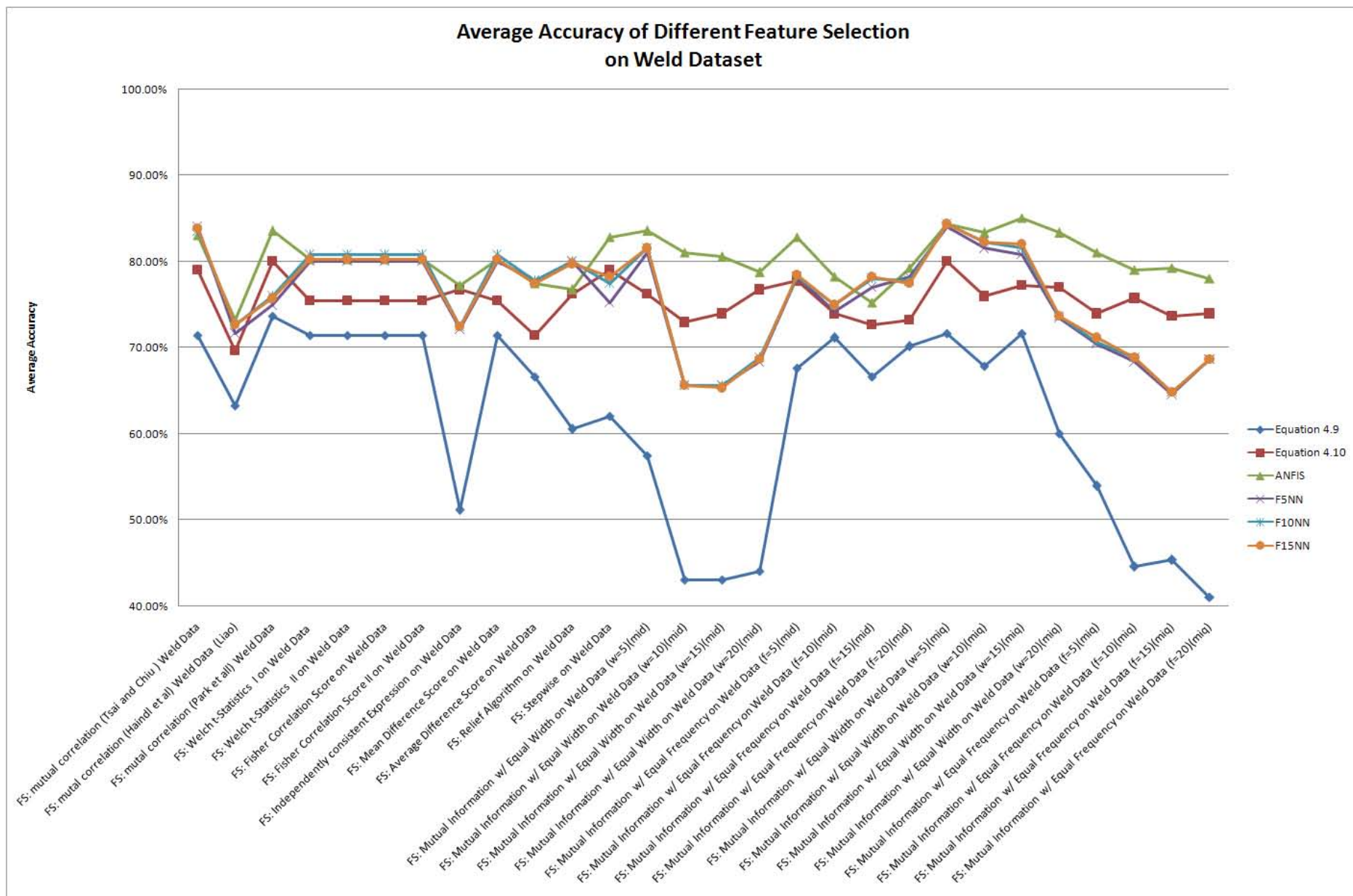


Figure 2. Average Accuracy of Different Feature Selection on Weld Dataset

Based on the average accuracy, the results of all feature selection-model combinations on Breast Cancer dataset are ranked, as shown in Table 38. The best result of 97.17% (in Table A7) was obtained by using the mutual correlation, as proposed by Tsai and Chiu [39], for feature selection and the Fuzzy 15NN model. The second and third best results were obtained by using the same mutual correlation algorithm for feature selection, and the proposed fuzzy model using Equation 4.10 and Fuzzy 5NN as their model respectively.

Table 38. Ranking of feature selection-model combination for the Breast Cancer dataset

Rank	Average Accuracy	Modeling Method	Feature Selection
1	97.17%	Fuzzy 15NN	FS: mutual correlation (Tsai and Chiu) on Breast Cancer Data
2	96.99%	Equation 4.10	FS: mutual correlation (Tsai and Chiu)
3	96.99%	Fuzzy 5NN	FS: mutual correlation (Tsai and Chiu)
4	96.46%	ANFIS	FS: mutual correlation (Tsai and Chiu)
5	96.46%	Fuzzy 10NN	FS: mutual correlation (Tsai and Chiu)
6	96.11%	Fuzzy 5NN	FS: Stepwise
7	96.11%	Fuzzy 10NN	FS: Stepwise
8	96.11%	Fuzzy 15NN	FS: Stepwise
9	95.95%	Fuzzy 5NN	FS: Mutual Information w/ Equal Frequency (f=10)(mid)
10	95.95%	Fuzzy 15NN	FS: Mutual Information w/ Equal Frequency (f=10)(mid)
11	95.70%	Fuzzy 10NN	FS: Mutual Information w/ Equal Frequency (f=10)(mid)
12	95.70%	Fuzzy 10NN	FS: Mutual Information w/ Equal Frequency (f=5)(miq)
13	95.70%	Fuzzy 15NN	FS: Mutual Information w/ Equal Frequency (f=5)(miq)
14	95.44%	Fuzzy 5NN	FS: Mutual Information w/ Equal Frequency (f=5)(miq)
15	95.44%	Fuzzy 15NN	FS: Mutual Information w/ Equal Frequency (f=5)(mid)
16	95.44%	Fuzzy 15NN	FS: Mutual Information w/ Equal Width (w=5)(miq)
17	95.22%	Fuzzy 15NN	FS: Relief Algorithm
18	95.19%	Fuzzy 10NN	FS: Mutual Information w/ Equal Frequency (f=5)(mid)
19	95.04%	Fuzzy 15NN	FS: Mutual correlation (Park et all)
20	94.94%	Fuzzy 5NN	FS: Mutual Information w/ Equal Frequency (f=5)(mid)

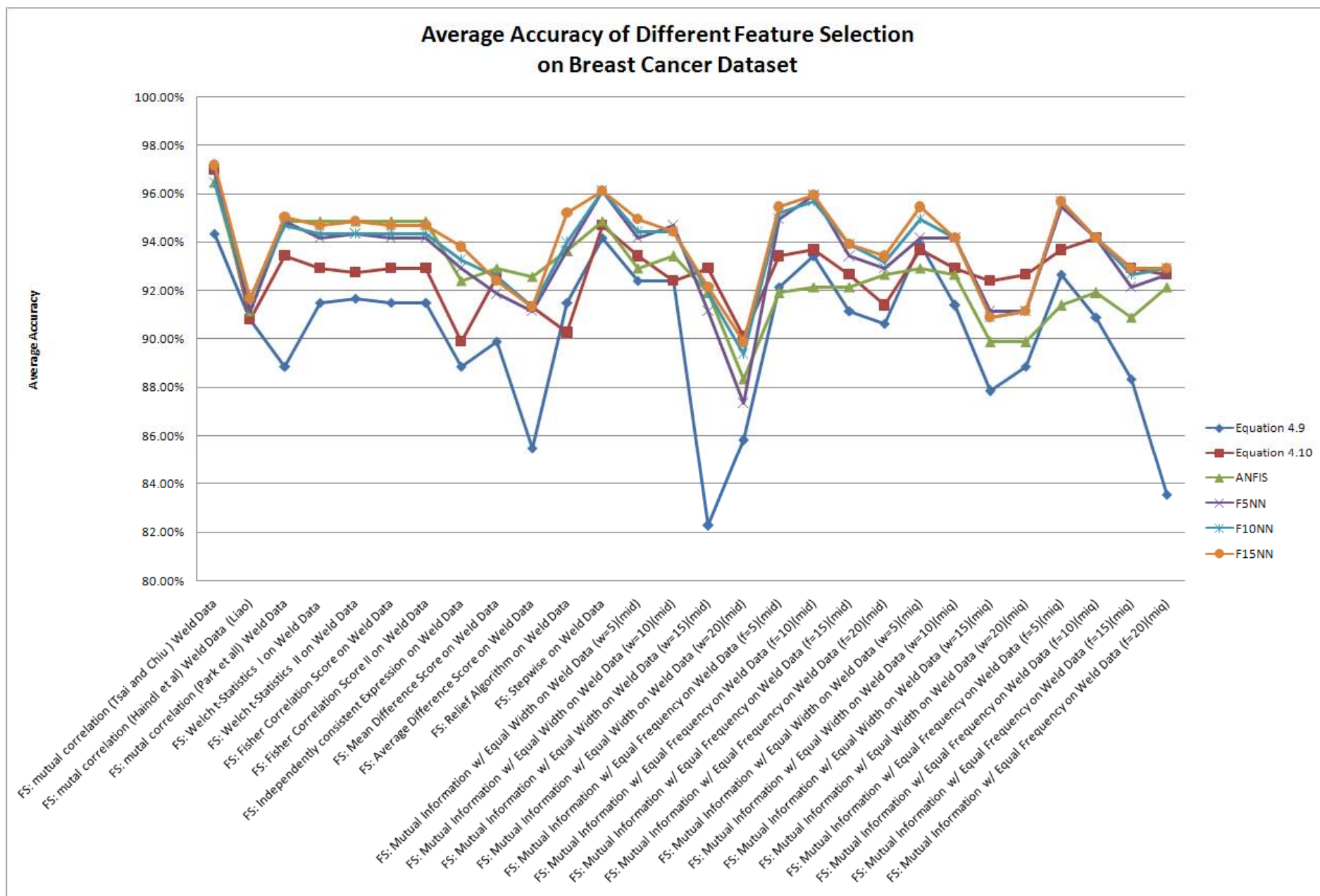


Figure 3. Average Accuracy of Different Feature Selection on Breast Cancer Dataset

Base on average accuracy, the results of all feature selection-model combinations on PIMA Diabetes dataset are ranked in Table 39. The three best results were obtained from ANFIS incorporating Welch's *t*-Statistics and Fisher's Correlation Score as the feature selection techniques, yielding average accuracies of 77.65%. The proposed fuzzy model employing Equation 4.10 yields an average accuracy of 77.39% using the stepwise algorithm as the preferred feature selection technique.

Table 39. Ranking of feature selection-model combination for the Breast Cancer dataset

Rank	Average Accuracy.	Model	Feature Selection
1	77.65%	ANFIS	FS: Welch t-Statistics I on PIMA Diabetes Data
2	77.65%	ANFIS	FS: Welch t-Statistics II on PIMA Diabetes Data
3	77.65%	ANFIS	FS: Fisher Correlation Score on PIMA Diabetes Data
4	77.65%	ANFIS	FS: Fisher Correlation Score II on PIMA Diabetes Data
5	77.39%	Equation 4.10	FS: Stepwise on PIMA Diabetes Data
6	76.86%	ANFIS	FS: Mutual correlation (Park et al) PIMA Diabetes Data
7	76.73%	Fuzzy 10NN	FS: Welch t-Statistics I on PIMA Diabetes Data
8	76.73%	Fuzzy 10NN	FS: Welch t-Statistics II on PIMA Diabetes Data
9	76.73%	Fuzzy 10NN	FS: Fisher Correlation Score on PIMA Diabetes Data
10	76.73%	Fuzzy 10NN	FS: Fisher Correlation Score II on PIMA Diabetes Data
11	76.73%	ANFIS	FS: Mean Difference Score on PIMA Diabetes Data
12	76.60%	Equation 4.10	FS: Mutual correlation (Tsai and Chiu) PIMA Diabetes Data
13	76.60%	Fuzzy 15NN	FS: Welch t-Statistics I on PIMA Diabetes Data
14	76.60%	Fuzzy 15NN	FS: Welch t-Statistics II on PIMA Diabetes Data
15	76.60%	Fuzzy 15NN	FS: Fisher Correlation Score on PIMA Diabetes Data
16	76.60%	Fuzzy 15NN	FS: Fisher Correlation Score II on PIMA Diabetes Data
17	76.21%	ANFIS	FS: Mutual correlation (Tsai and Chiu) PIMA Diabetes Data
18	76.08%	ANFIS	FS: Stepwise on PIMA Diabetes Data
19	75.82%	Fuzzy 15NN	FS: Mutual correlation (Park et al) PIMA Diabetes Data
20	75.82%	Fuzzy 15NN	FS: Mean Difference Score on PIMA Diabetes Data

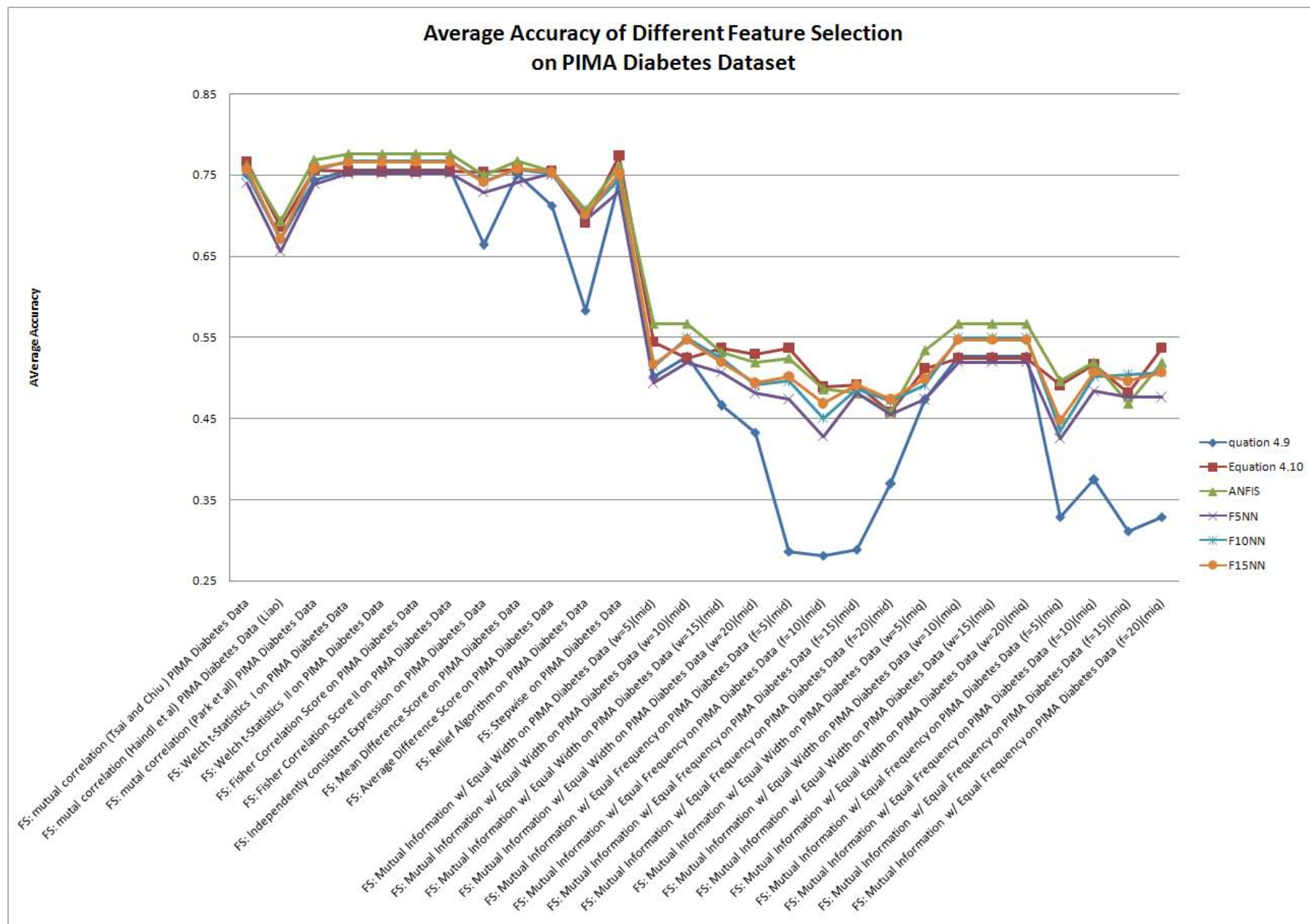


Figure 4. Average Accuracy of Different Feature Selection on PIMA Diabetes Dataset

6. Conclusion

In this research a fuzzy modeling and reasoning method has been clearly defined, implemented and the accuracies of models with different combinations of feature subsets of size 3 have been obtained. The acquired results are then compared with the results generated by ANFIS, fuzzy K-nearest neighbor with three different numbers of neighbors, i.e. five, ten and fifteen neighbors. ANFIS and fuzzy K-nearest neighbor are two well known fuzzy algorithms/models. All these models are applied to three different datasets as shown in Table 1. Then the computational time of each model on different datasets is measured and reported. It is concluded that fuzzy K-nearest neighbor algorithm is the fastest, ANFIS is the slowest and our fuzzy modeling and reasoning method is in between of the two. It is shown that ANFIS is the most accurate of all tested algorithms/models and our proposed fuzzy method incorporating Equation 4.10 is comparable with fuzzy k-nearest neighbors.

Nine different feature selection techniques are proposed and implemented. All of these nine methods are from the filter category of feature selection methods. These nine methods are mutual correlation, mutual information, stepwise algorithm, Welch's *t*-statistics, Fisher's correlation score, independently consistent expression, mean difference score, average difference score, and relief algorithm. Note that variations of four of these methods were also implemented. Specifically, they are three variations of mutual correlation, 16 variations of mutual information, two variations of Welch's *t*-Statistics and two variations of Fisher's correlation score. Among the feature selection technique implemented in this thesis, mutual correlation, stepwise algorithm, and mutual information using equal width discretization technique with medium numbers of intervals or discretized values, (=15) provided the highest accuracies. The accuracies of the proposed fuzzy model, as well as other models, can be improved if there are more data records available. In this research only three features were considered because by employing the model using only these three features in the worst case scenario $5 \times 5 \times 5 = 250$ rules need

to be generated to train the system and at least five data records for each rule is required. Therefore, if the number of features is increased, assuming by only one feature or using four features instead of three, the model size is increased five times requiring $5 \times 5 \times 5 \times 5 \times 2 = 1250$ rules, and five times more data records are needed. If the number of records is increased, the time to train the system is also going to rise, and using ANFIS it is going to take a while to train the system, and the same arguments applies to fuzzy K-nearest neighbor. Fuzzy K-nearest does not generate any rules, instead searches within its neighbors, therefore, the number of points it has to go through is $(\text{number of neighbors}^{\text{number of data records}})$. Hence, as the number of data records is increased, the computational time of fuzzy K-nearest neighbor will also increase by factor of its neighbors per data point.

It is shown that the max operator used in Equation 4.9 as an aggregation operator is not very effective when the number of data is limited, and it cannot obtain the best value as its rule weight resulting in its poor performance. It is recommended to find better aggregation operator to replace the max operator in Equation 4.9, in order to obtain better results.

For future research, different evaluation criteria such as area under the curve (AUC) can be used instead of accuracies to measure the performance of the proposed model. Also, there are different statistical approaches to select the optimal number of clusters, which can be applied in future works.

References

- [1] T. W. Liao and J. W. Ni, "An automated radiographic NDT system for weld inspection .1. Weld extraction," *Ndt & E International*, vol. 29, pp. 157-162, Jun 1996.
- [2] C. Hayes, "The ABC's of nondestructive weld examination," *Welding Journal*, vol. 76, pp. 46-51, May 1997.
- [3] T. W. Liao, D. M. Li, and Y. M. Li, "Extraction of welds from radiographic images using fuzzy classifiers," *Information Sciences*, vol. 126, pp. 21-40, Jul 2000.
- [4] T. W. Liao, A. K. Celmins, and R. J. Hammell, "A fuzzy c-means variant for the generation of fuzzy term sets," *Fuzzy Sets and Systems*, vol. 135, pp. 241-257, Apr 2003.
- [5] W. Daum, P. Rose, H. Heidt, and J. H. Builtjes, "AUTOMATIC RECOGNITION OF WELD DEFECTS IN X-RAY INSPECTION," *British Journal of Non-Destructive Testing*, vol. 29, pp. 79-82, Mar 1987.
- [6] A. Gayer, A. Saya, and A. Shiloh, "Automatic Recognition of Welding Defects in Real-Time Radiography," *Ndt International*, vol. 23, pp. 131-136, Jun 1990.
- [7] K. Murakami, "Image processing for non-destructive testing," *Weld International*, vol. 4, pp. 144-149, 1990.
- [8] O. T. Kato Y, Matsui S, Itoga K, Harada T, Sugimoto K, Michiba K, Iuchi S, Kawano S., "Development of an automatic weld defect identification system for radiographic testing," *Welding Word*, vol. 30, pp. 182-188, 1992.
- [9] R. Hyatt, G. E. Kechter, and S. Nagashima, "A method for defect segmentation in digital radiographs of pipeline girth welds," *Materials Evaluation*, vol. 54, pp. 925-928, Aug 1996.
- [10] T. W. Liao and K. Tang, "Automated extraction of welds from digitized radiographic images based on MLP neural networks," *Applied Artificial Intelligence*, vol. 11, pp. 197-218, Apr-May 1997.
- [11] T. W. Liao and Y. M. Li, "An automated radiographic NDT system for weld inspection: Part II - Flaw detection," *Ndt & E International*, vol. 31, pp. 183-192, Jun 1998.

- [12] S. Y. Aoki K, "Application of artificial neural network to discrimination of defect type automatic radiographic testing of welds," *ISI International*, vol. 39, pp. 1081-1087, 1999.
- [13] T. W. Liao, D. M. Li, and Y. M. Li, "Detection of welding flaws from radiographic images with fuzzy clustering methods," *Fuzzy Sets and Systems*, vol. 108, pp. 145-158, Dec 1999.
- [14] G. Wang and T. W. Liao, "Automatic identification of different types of welding defects in radiographic images," *Ndt & E International*, vol. 35, pp. 519-528, Dec 2002.
- [15] T. W. Liao, "Classification of welding flaw types with fuzzy expert systems," *Expert Systems with Applications*, vol. 25, pp. 101-111, Jul 2003.
- [16] L.-X. Wang, Mendel, J. M., "Generating fuzzy rules by learning from examples," *IEEE Transaction on System, Man, and Cybernetics*, vol. 22, pp. 1414- 1427, 1992.
- [17] V. Kaftandjian, O. Dupuis, D. Babot, and Y. M. Zhu, "Uncertainty modelling using Dempster-Shafer theory for improving detection of weld defects," *Pattern Recognition Letters*, vol. 24, pp. 547-564, Jan 2003.
- [18] T. W. Liao, "Fuzzy reasoning based automatic inspection of radiographic welds: weld recognition," *Journal of Intelligent Manufacturing*, vol. 15, pp. 69-85, Feb 2004.
- [19] X. Wang and B. S. Wong, "Segmentation of radiographic images using fuzzy c-means algorithm," *Insight*, vol. 47, pp. 631-633, Oct 2005.
- [20] M. K. Felisberto, H. S. Lopes, T. M. Centeno, and L. V. de Arruda, "An object detection and recognition system for weld bead extraction from digital radiographs," *Computer Vision and Image Understanding*, vol. 102, pp. 238-249, Jun 2006.
- [21] R. R. da Silva, L. P. Caloba, M. H. S. Siqueira, and J. M. A. Rebello, "Pattern recognition of weld defects detected by radiographic test," *Ndt & E International*, vol. 37, pp. 461-470, Sep 2004.
- [22] T. W. Liao, "Improving the Accuracy of Computer-Aided Radiographic Weld Inspection by Feature Selection," *NDT & E International*, vol. (Submitted), 2007.
- [23] L. A. Zadeh, "FUZZY SETS," *Information and Control*, vol. 8, pp. 338-&, 1965.

- [24] L. A. Zadeh, "OUTLINE OF A NEW APPROACH TO ANALYSIS OF COMPLEX SYSTEMS AND DECISION PROCESSES," *Ieee Transactions on Systems Man and Cybernetics*, vol. SMC3, pp. 28-44, 1973.
- [25] E. H. Mamdani and S. Assilian, "EXPERIMENT IN LINGUISTIC SYNTHESIS WITH A FUZZY LOGIC CONTROLLER," *International Journal of Man-Machine Studies*, vol. 7, pp. 1-13, 1975.
- [26] T. Takagi and M. Sugeno, "Fuzzy Identification of Systems and Its Applications to Modeling and Control," *Ieee Transactions on Systems Man and Cybernetics*, vol. 15, pp. 116-132, 1985.
- [27] M. Sugeno and G. T. Kang, "FUZZY MODELING AND CONTROL OF MULTILAYER INCINERATOR," *Fuzzy Sets and Systems*, vol. 18, pp. 329-345, Apr 1986.
- [28] S. Guillaume, "Designing fuzzy inference systems from data: An interpretability-oriented review," *Ieee Transactions on Fuzzy Systems*, vol. 9, pp. 426-443, Jun 2001.
- [29] T. W. Liao, "Mining Human Interpretable Knowledge Using Automatic Data-Driven Fuzzy Modeling Methods – A Review," *Chapter 15 in Data Mining and Knowledge Discovery Approaches Based on Rule Induction Techniques*, E. Triantaphyllou and G. Felici (Eds.), Kluwer Acad. Publishers, Massive Computing Series, pp. 495-550, 2006.
- [30] Y. H. Shi, R. Eberhart, and Y. B. Chen, "Implementation of evolutionary fuzzy systems," *Ieee Transactions on Fuzzy Systems*, vol. 7, pp. 109-119, Apr 1999.
- [31] M. Sugeno and T. Yasukawa, "A fuzzy logic based approach to qualitative modeling," *IEEE Trans. Fuzzy Systems*, vol. 1, pp. 7-31, 1993.
- [32] J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," *J. Cybernet*, vol. 3, pp. 32-57, 1974.
- [33] J. C. Bezdek, "Pattern Recognition with Fuzzy Objective Function Algorithms," *Plenum Press, New York and London*, 1987.
- [34] S. Medasani, J. Kim, and R. Krishnapuram, "An overview of membership function generation techniques for pattern recognition," *International Journal of Approximate Reasoning*, vol. 19, pp. 391-417, Oct-Nov 1998.
- [35] K. Josien and T. W. Liao, "Simultaneous grouping of parts and machines with an integrated fuzzy clustering method," *Fuzzy Sets and Systems*, vol. 126, pp. 1-21, Feb 2002.

- [36] M. Delgado, A. F. GomezSkarmeta, and F. Martin, "A fuzzy clustering-based rapid prototyping for fuzzy rule-based modeling," *Ieee Transactions on Fuzzy Systems*, vol. 5, pp. 223-233, May 1997.
- [37] M. Haindl, P. Somol, D. Ververidis, and C. Kotropoulos, "Feature selection based on mutual correlation," in *Progress in Pattern Recognition, Image Analysis and Applications, Proceedings*. vol. 4225, 2006, pp. 569-577.
- [38] J. S. Park, K. M. Shazzad, and D. S. Kim, "Toward modeling lightweight intrusion detection system through correlation-based hybrid feature selection," in *Information Security and Cryptology, Proceedings*. vol. 3822, 2005, pp. 279-289.
- [39] C. Y. Tsai and C. C. Chiu, "A case-based reasoning system for PCB principal process parameter identification," *Expert Systems with Applications*, vol. 32, pp. 1183-1193, May 2007.
- [40] J. R. Quinlan, "C4.5: Programs for Machine Learning.," *Morgan Kaufmann*, 1993.
- [41] H. C. Peng, F. H. Long, and C. Ding, "Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy," *Ieee Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, pp. 1226-1238, Aug 2005.
- [42] C. Ding and H. C. Peng, "Minimum redundancy feature selection from microarray gene expression data," *Journal of Bioinformatics and Computational Biology*, vol. 3, pp. 185-205, 2005.
- [43] R. F. Gunst and R. L. Mason, "Regression Analysis and Its Application," *Marcel Dekker Inc. , New York and Basel*, 1980.
- [44] L. H. Loo, S. Roberts, L. Hrebien, and M. Kam, "New criteria for selecting differentially expressed genes - Filter-based feature selection techniques for better detection of changes in the distributions of expression levels," *Ieee Engineering in Medicine and Biology Magazine*, vol. 26, pp. 17-26, Mar-Apr 2007.
- [45] R. Bijlani, Y. H. Cheng, D. A. Pearce, A. I. Brooks, and M. Ogihara, "Prediction of biologically significant components from microarray data: Independently Consistent Expression Discriminator (ICED)," *Bioinformatics*, vol. 19, pp. 62-70, Jan 2003.
- [46] K. Kira and L. A. Rendell, "A practical approach to feature selection," *Proceedings of the Ninth International Conference on Machine Learning (ICML-92), Aberdeen, Scotland, Morgan Kaufmann, CA*, pp. 129-134, 1992.

- [47] J. S. R. Jang, "Anfis - Adaptive-Network-Based Fuzzy Inference System," *Ieee Transactions on Systems Man and Cybernetics*, vol. 23, pp. 665-685, May-Jun 1993.
- [48] J. M. Keller, M. R. Gray, and J. A. Givens, "A FUZZY K-NEAREST NEIGHBOR ALGORITHM," *Ieee Transactions on Systems Man and Cybernetics*, vol. 15, pp. 580-585, 1985.

Appendix

A. Results of Breast Cancer Dataset

Table A 1. Average errors of five folds on Breast Cancer data from EQ. 4.9

i	j	k	Average Error Bases on Equation (4.9)	Standard Deviation
2	24	28	4	1.4142
9	24	28	4.2	1.0954
23	24	28	4.4	0.8944
22	24	28	4.4	1.1402
24	26	28	4.4	1.5166
7	22	23	4.4	2.0736
23	24	25	4.4	2.3022
8	24	28	4.6	1.5166
14	24	28	4.6	1.5166
17	24	28	4.6	1.5166
24	27	28	4.6	1.5166
7	21	26	4.6	2.1909
22	24	29	4.6	2.4083
8	21	22	4.6	2.7019
24	25	28	4.8	0.8367
6	22	24	4.8	1.3038
10	24	28	4.8	1.3038
8	22	24	4.8	1.7889
8	24	25	4.8	2.1679
19	24	28	5	1.4142
20	24	28	5	1.4142
3	21	25	5	1.5811
5	24	28	5	1.5811
16	24	28	5	1.5811
1	25	27	5	1.7321
24	25	27	5	2.5495
24	28	30	5.2	1.3038
7	21	25	5.2	1.6432
2	23	25	5.2	1.7889
13	24	28	5.2	1.7889
3	7	21	5.2	1.9235
24	28	29	5.2	1.9235
2	21	25	5.2	2.5884
5	24	29	5.2	2.5884
2	7	24	5.2	2.7749
7	21	22	5.2	2.8636
8	24	26	5.4	0.8944
1	21	25	5.4	1.5166
8	20	24	5.4	1.5166
21	25	27	5.4	1.5166
2	24	30	5.4	1.6733
24	26	29	5.4	1.9494
7	24	28	5.4	2.0736
5	24	27	5.4	2.1909
7	21	24	5.4	2.3022
8	23	24	5.6	0.8944
3	25	27	5.6	1.5166
21	22	27	5.6	1.5166
18	24	28	5.6	1.6733

Table A 2. Average errors of five folds on Breast Cancer data from EQ. 4.10

i	j	k	Average Error Bases on Equation (4.10)	Standard Deviation
2	24	28	4	1.4142
9	24	28	4.2	1.0954
23	24	28	4.4	0.8944
22	24	28	4.4	1.1402
24	26	28	4.4	1.5166
7	22	23	4.4	2.0736
23	24	25	4.4	2.3022
8	24	28	4.6	1.5166
14	24	28	4.6	1.5166
17	24	28	4.6	1.5166
24	27	28	4.6	1.5166
7	21	26	4.6	2.1909
22	24	29	4.6	2.4083
8	21	22	4.6	2.7019
24	25	28	4.8	0.8367
6	22	24	4.8	1.3038
10	24	28	4.8	1.3038
8	22	24	4.8	1.7889
8	24	25	4.8	2.1679
19	24	28	5	1.4142
20	24	28	5	1.4142
3	21	25	5	1.5811
5	24	28	5	1.5811
16	24	28	5	1.5811
1	25	27	5	1.7321
24	25	27	5	2.5495
24	28	30	5.2	1.3038
7	21	25	5.2	1.6432
2	23	25	5.2	1.7889
13	24	28	5.2	1.7889
3	7	21	5.2	1.9235
24	28	29	5.2	1.9235
2	21	25	5.2	2.5884
5	24	29	5.2	2.5884
2	7	24	5.2	2.7749
7	21	22	5.2	2.8636
8	24	26	5.4	0.8944
1	21	25	5.4	1.5166
8	20	24	5.4	1.5166
21	25	27	5.4	1.5166
2	24	30	5.4	1.6733
24	26	29	5.4	1.9494
7	24	28	5.4	2.0736
5	24	27	5.4	2.1909
7	21	24	5.4	2.3022
8	23	24	5.6	0.8944
3	25	27	5.6	1.5166
21	22	27	5.6	1.5166
18	24	28	5.6	1.6733
8	24	29	5.6	1.8166

Table A 3. Average errors of five folds on Breast Cancer data from ANFIS

i	j	k	Average Error based on ANFIS	Standard Deviation
21	22	28	3.8	1.4832
1	22	28	4	0.7071
22	24	28	4	1.5811
1	24	25	4.2	1.0954
3	24	28	4.4	1.5166
21	23	25	4.4	1.5166
1	24	28	4.4	1.9494
1	8	24	4.6	0.5477
3	22	28	4.6	0.8944
1	7	24	4.6	1.1402
3	7	24	4.6	1.1402
22	23	28	4.6	1.3416
24	27	28	4.6	2.6077
5	7	21	4.6	3.0496
1	21	28	4.8	0.8367
3	24	25	4.8	1.4832
2	24	27	4.8	1.7889
1	24	26	4.8	1.9235
5	24	27	4.8	1.9235
5	21	27	4.8	2.7749
1	7	21	5	1
23	25	28	5	1
5	23	27	5	1.2247
21	23	28	5	1.2247
24	28	29	5	1.2247
2	24	28	5	1.8708
23	24	28	5	1.8708
3	7	23	5	2
3	24	26	5	2
4	24	25	5	2.3452
8	22	24	5	2.3452
21	24	25	5	2.3452
5	21	22	5	2.8284
5	24	28	5.2	0.8367
2	23	28	5.2	1.0954
3	7	21	5.2	1.0954
7	22	24	5.2	1.0954
21	24	28	5.2	1.0954
3	21	28	5.2	1.3038
4	21	28	5.2	1.3038
2	3	28	5.2	1.4832
3	23	28	5.2	1.4832
15	24	28	5.2	1.4832
17	24	26	5.2	1.4832
1	23	25	5.2	1.6432
5	21	28	5.2	1.6432
10	21	28	5.2	1.6432
15	23	28	5.2	1.6432
23	24	25	5.2	1.6432
7	22	23	5.2	1.7889

Table A 4. Average errors of five folds on Breast Cancer data from Fuzzy SNN

i	j	k	Average Error based on Fuzzy KNN (N=5)	Standard Deviation
22	24	28	2.6	0.8944
21	22	28	2.8	1.3038
8	21	22	3.2	1.4832
8	22	23	3.2	1.4832
8	22	24	3.2	1.4832
22	23	28	3.2	1.4832
2	23	25	3.4	2.1909
3	21	28	3.6	1.1402
2	23	28	3.6	1.5166
21	22	25	3.6	1.5166
4	22	28	3.8	1.3038
18	23	25	3.8	1.3038
2	21	28	3.8	1.4832
22	23	25	3.8	1.4832
2	21	25	3.8	1.9235
4	22	29	4	2
7	16	23	4	2.5495
1	22	28	4.2	1.3038
14	22	28	4.2	1.3038
2	24	28	4.2	1.6432
3	22	28	4.2	1.6432
5	21	27	4.2	1.6432
7	21	22	4.2	1.6432
3	22	29	4.2	1.7889
18	21	25	4.2	1.9235
2	24	25	4.2	2.1679
7	8	24	4.2	2.2804
7	16	24	4.2	2.2804
22	24	29	4.2	2.5884
1	21	28	4.4	0.8944
3	24	28	4.4	0.8944
4	21	28	4.4	0.8944
11	23	25	4.4	1.1402
16	21	28	4.4	1.1402
4	23	25	4.4	1.5166
21	25	26	4.4	1.6733
1	23	25	4.4	2.0736
5	24	27	4.4	2.0736
20	23	25	4.4	2.0736
15	21	25	4.4	2.3022
7	16	21	4.4	2.51
21	25	27	4.4	3.4351
1	28	30	4.6	1.1402
13	24	28	4.6	1.1402
13	21	28	4.6	1.3416
12	23	25	4.6	1.5166
17	21	30	4.6	1.5166
14	21	28	4.6	1.8166
1	22	29	4.6	2.0736
7	8	21	4.6	2.0736

Table A 5. Average errors of five folds on Breast Cancer data from Fuzzy 10NN

i	j	k	Average Error based on Fuzzy KNN (N=10)	Standard Deviation
22	24	28	2.6	0.8944
21	22	28	2.8	1.0954
8	21	22	2.8	1.9235
21	22	25	3	1
22	23	28	3	1.2247
8	22	24	3	1.5811
2	21	25	3	1.7321
8	22	23	3	2
2	23	25	3	2.3452
21	25	27	3.2	2.7749
2	23	28	3.4	0.5477
2	21	28	3.4	1.1402
4	22	28	3.4	1.1402
5	24	27	3.4	1.5166
2	24	28	3.4	1.8166
3	21	28	3.6	1.1402
18	23	25	3.6	1.1402
5	21	27	3.6	1.5166
19	23	25	3.6	2.4083
18	21	25	3.8	0.8367
22	24	25	3.8	0.8367
21	22	27	3.8	2.1679
7	16	24	3.8	2.49
21	22	29	3.8	2.5884
1	21	28	4	1
11	23	25	4	1.4142
4	22	29	4	1.5811
14	22	28	4	1.5811
22	23	25	4	1.5811
13	23	25	4	1.7321
1	22	28	4	1.8708
5	23	27	4	2
22	24	29	4	2.9155
22	23	29	4	3.0822
3	24	28	4.2	0.8367
4	24	28	4.2	0.8367
16	21	28	4.2	0.8367
4	21	28	4.2	1.0954
12	23	25	4.2	1.4832
4	23	25	4.2	1.6432
4	24	25	4.2	1.7889
14	23	25	4.2	1.7889
3	22	28	4.2	1.9235
4	25	27	4.2	1.9235
7	21	22	4.2	1.9235
1	23	25	4.2	2.1679
20	23	25	4.2	2.1679
22	23	27	4.2	2.2804
23	24	25	4.2	2.7749
19	23	30	4.4	0.5477

Table A 6. Average errors of five folds on Breast Cancer data from Fuzzy 15NN

i	j	k	Average Error based on Fuzzy KNN (N=15)	Standard Deviation
22	24	28	2.6	0.8944
2	24	28	2.8	1.3038
21	22	28	2.8	1.4832
2	21	25	2.8	1.9235
8	21	22	2.8	1.9235
8	22	23	2.8	1.9235
22	23	28	3	1.2247
2	23	25	3	2.3452
21	22	25	3.2	0.8367
8	22	24	3.2	1.7889
2	23	28	3.4	0.5477
1	22	28	3.4	1.1402
2	21	28	3.4	1.1402
4	22	28	3.4	1.1402
3	21	28	3.6	1.1402
5	24	27	3.6	1.5166
19	23	25	3.6	2.4083
21	25	27	3.6	3.1305
3	24	28	3.8	0.8367
14	22	28	3.8	0.8367
22	24	25	3.8	0.8367
1	21	28	3.8	1.3038
22	23	25	3.8	1.3038
3	22	29	3.8	1.6432
3	22	28	3.8	1.7889
7	21	22	3.8	1.7889
5	21	27	4	1.4142
12	23	25	4	1.4142
7	22	24	4	1.5811
4	23	25	4	1.7321
2	24	25	4	1.8708
5	22	23	4	2
11	23	25	4	2
21	22	27	4	2
1	23	25	4	2.3452
15	21	25	4	2.3452
9	22	23	4	2.5495
21	22	29	4	2.8284
22	24	29	4	2.9155
19	23	30	4.2	0.4472
2	9	21	4.2	0.8367
4	21	28	4.2	1.0954
4	24	28	4.2	1.0954
2	24	27	4.2	1.3038
4	22	29	4.2	1.4832
7	22	23	4.2	1.4832
12	21	25	4.2	1.4832
15	23	25	4.2	1.6432
13	23	25	4.2	1.7889
14	23	25	4.2	1.7889

Table A 7. Feature Selection with Mutual Correlation on Breast Cancer Data

FS: mutual correlation (Tsai and Chiu) Breast Cancer Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	28	21	22	7	5	4	2	3	3	93.81%	95.58%	96.46%	98.23%	97.35%	97.35%
Fold2	28	21	22	4	4	6	4	4	5	96.46%	96.46%	94.69%	96.46%	96.46%	95.58%
Fold3	28	21	22	9	2	2	4	3	2	92.04%	98.23%	98.23%	96.46%	97.35%	98.23%
Fold4	28	21	2	6	3	5	4	7	3	94.69%	97.35%	95.58%	96.46%	93.81%	97.35%
Fold5	28	21	22	6	3	3	3	3	3	94.69%	97.35%	97.35%	97.35%	97.35%	97.35%
Average Error										94.34%	96.99%	96.46%	96.99%	96.46%	97.17%

Table A 8. Feature Selection with Mutual Correlation on Breast Cancer Data

FS: mutal correlation (Haindl et al) Breast Cancer Data (Liao)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	1	5	29	11	13	11	9	7	8	90.27%	88.50%	90.27%	92.04%	93.81%	92.92%
Fold2	1	15	29	11	13	7	8	8	8	90.27%	88.50%	93.81%	92.92%	92.92%	92.92%
Fold3	1	15	29	10	10	8	10	10	10	91.15%	91.15%	92.92%	91.15%	91.15%	91.15%
Fold4	1	15	29	6	3	9	8	8	7	94.69%	97.35%	92.04%	92.92%	92.92%	93.81%
Fold5	1	15	29	14	13	15	15	14	14	87.61%	88.50%	86.73%	86.73%	87.61%	87.61%
Average Error										90.80%	90.80%	91.15%	91.15%	91.68%	91.68%

Table A 9. Feature Selection with Mutual Correlation on Breast Cancer Data

FS: mutal correlation (Park et al) Breast Cancer Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	8	21	28	30	4	6	5	6	5	73.45%	96.46%	94.69%	95.58%	94.69%	95.58%
Fold2	8	21	28	6	7	8	7	7	7	94.69%	93.81%	92.92%	93.81%	93.81%	93.81%
Fold3	8	21	28	13	16	5	7	7	6	88.50%	85.84%	95.58%	93.81%	93.81%	94.69%
Fold4	21	23	28	4	6	6	5	5	5	96.46%	94.69%	94.69%	95.58%	95.58%	95.58%
Fold5	8	21	28	10	4	4	5	5	5	91.15%	96.46%	96.46%	95.58%	95.58%	95.58%
Average Error										88.85%	93.45%	94.87%	94.87%	94.69%	95.04%

Table A 10. Feature Selection with Mutual Information on Breast Cancer Data

FS: Mutual Information w/ Equal Width on breast cancer Data (w=5)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	28	22	21	7	5	4	2	3	3	91.14%	93.67%	94.94%	97.47%	96.20%	96.20%
Fold2	28	24	2	4	8	6	6	6	4	94.94%	89.87%	92.41%	92.41%	92.41%	94.94%
Fold3	8	2	24	8	7	8	9	7	7	89.87%	91.14%	89.87%	88.61%	91.14%	91.14%
Fold4	23	2	28	6	2	4	3	3	3	92.41%	97.47%	94.94%	96.20%	96.20%	96.20%
Fold5	8	22	21	5	4	6	3	3	3	93.67%	94.94%	92.41%	96.20%	96.20%	96.20%
Average Error										92.41%	93.42%	92.91%	94.18%	94.43%	94.94%

Table A 11. Feature Selection with Mutual Information on Breast Cancer Data

FS: Mutual Information w/ Equal Width on breast cancer Data (w=10)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	28	22	24	4	3	3	3	3	3	94.94%	96.20%	96.20%	96.20%	96.20%	96.20%
Fold2	28	11	24	7	9	6	4	5	5	91.14%	88.61%	92.41%	94.94%	93.67%	93.67%
Fold3	28	11	22	8	8	7	6	6	6	89.87%	89.87%	91.14%	92.41%	92.41%	92.41%
Fold4	23	17	28	5	6	6	5	5	5	93.67%	92.41%	92.41%	93.67%	93.67%	93.67%
Fold5	23	22	28	6	4	4	3	3	3	92.41%	94.94%	94.94%	96.20%	96.20%	96.20%
Average Error										92.41%	92.41%	93.42%	94.68%	94.43%	94.43%

Table A 12. Feature Selection with Mutual Information on Breast Cancer Data

FS: Mutual Information w/ Equal Width on breast cancer Data (w=15)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	23	20	8	11	6	5	8	8	7	86.08%	92.41%	93.67%	89.87%	89.87%	91.14%
Fold2	23	17	28	17	7	8	8	6	7	78.48%	91.14%	89.87%	89.87%	92.41%	91.14%
Fold3	23	20	28	11	5	6	7	6	5	86.08%	93.67%	92.41%	91.14%	92.41%	93.67%
Fold4	23	17	2	13	5	7	5	5	5	83.54%	93.67%	91.14%	93.67%	93.67%	93.67%
Fold5	23	20	28	18	5	6	7	7	7	77.22%	93.67%	92.41%	91.14%	91.14%	91.14%
Average Error										82.28%	92.91%	91.90%	91.14%	91.90%	92.15%

Table A 13. Feature Selection with Mutual Information on Breast Cancer Data

FS: Mutual Information w/ Equal Width on breast cancer Data (w=20)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	23	17	14	14	10	8	11	10	10	82.28%	87.34%	89.87%	86.08%	87.34%	87.34%
Fold2	23	17	14	8	8	10	13	10	10	89.87%	89.87%	87.34%	83.54%	87.34%	87.34%
Fold3	23	17	14	15	4	10	8	8	7	81.01%	94.94%	87.34%	89.87%	89.87%	91.14%
Fold4	23	17	14	11	9	8	8	6	5	86.08%	88.61%	89.87%	89.87%	92.41%	93.67%
Fold5	23	17	14	8	8	10	10	8	8	89.87%	89.87%	87.34%	87.34%	89.87%	89.87%
Average Error										85.82%	90.13%	88.35%	87.34%	89.37%	89.87%

Table A 14. Feature Selection with Mutual Information on Breast Cancer Data

FS: Mutual Information w/ Equal Frequency on breast cancer Data (f=5)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	23	22	28	8	5	6	3	3	3	89.87%	93.67%	92.41%	96.20%	96.20%	96.20%
Fold2	23	25	2	7	8	11	7	7	7	91.14%	89.87%	86.08%	91.14%	91.14%	91.14%
Fold3	8	22	21	6	5	7	4	4	4	92.41%	93.67%	91.14%	94.94%	94.94%	94.94%
Fold4	23	22	28	3	4	4	1	1	1	96.20%	94.94%	94.94%	98.73%	98.73%	98.73%
Fold5	23	25	8	7	4	4	5	4	3	91.14%	94.94%	94.94%	93.67%	94.94%	96.20%
Average Error										92.15%	93.42%	91.90%	94.94%	95.19%	95.44%

Table A 15. Feature Selection with Mutual Information on Breast Cancer Data

FS: Mutual Information w/ Equal Frequency on breast cancer Data (f=10)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	23	22	28	8	5	6	3	3	3	89.87%	93.67%	92.41%	96.20%	96.20%	96.20%
Fold2	23	22	8	7	8	7	5	5	5	91.14%	89.87%	91.14%	93.67%	93.67%	93.67%
Fold3	23	22	8	4	4	7	4	4	4	94.94%	94.94%	91.14%	94.94%	94.94%	94.94%
Fold4	23	22	28	3	4	4	1	1	1	96.20%	94.94%	94.94%	98.73%	98.73%	98.73%
Fold5	23	22	8	4	4	7	3	4	3	94.94%	94.94%	91.14%	96.20%	94.94%	96.20%
Average Error										93.42%	93.67%	92.15%	95.95%	95.70%	95.95%

Table A 16. Feature Selection with Mutual Information on Breast Cancer Data

FS: Mutual Information w/ Equal Frequency on breast cancer Data (f=15)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	23	25	8	7	3	4	4	2	2	91.14%	96.20%	94.94%	94.94%	97.47%	97.47%
Fold2	23	25	8	10	9	9	8	8	8	87.34%	88.61%	88.61%	89.87%	89.87%	89.87%
Fold3	23	25	8	8	8	7	6	6	6	89.87%	89.87%	91.14%	92.41%	92.41%	92.41%
Fold4	23	25	28	5	5	4	3	4	4	93.67%	93.67%	94.94%	96.20%	94.94%	94.94%
Fold5	23	2	8	5	4	7	5	4	4	93.67%	94.94%	91.14%	93.67%	94.94%	94.94%
Average Error										91.14%	92.66%	92.15%	93.42%	93.92%	93.92%

Table A 17. Feature Selection with Mutual Information on Breast Cancer Data

FS: Mutual Information w/ Equal Frequency on breast cancer Data (f=20)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	23	29	28	5	9	6	4	5	5	93.67%	88.61%	92.41%	94.94%	93.67%	93.67%
Fold2	23	25	28	7	7	6	8	7	7	91.14%	91.14%	92.41%	89.87%	91.14%	91.14%
Fold3	23	25	8	8	8	7	6	6	6	89.87%	89.87%	91.14%	92.41%	92.41%	92.41%
Fold4	23	25	8	10	6	6	5	5	5	87.34%	92.41%	92.41%	93.67%	93.67%	93.67%
Fold5	23	22	8	7	4	4	5	4	3	91.14%	94.94%	94.94%	93.67%	94.94%	96.20%
Average Error										90.63%	91.39%	92.66%	92.91%	93.16%	93.42%

Table A 18. Feature Selection with Mutual Information on Breast Cancer Data

FS: Mutual Information w/ Equal Width on breast cancer Data (w=5)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	28	22	24	4	3	3	3	3	3	94.94%	96.20%	96.20%	96.20%	96.20%	96.20%
Fold2	28	2	24	4	8	6	6	6	4	94.94%	89.87%	92.41%	92.41%	92.41%	94.94%
Fold3	8	2	24	8	7	8	9	7	7	89.87%	91.14%	89.87%	88.61%	91.14%	91.14%
Fold4	23	25	2	3	3	5	2	1	1	96.20%	96.20%	93.67%	97.47%	98.73%	98.73%
Fold5	8	22	21	5	4	6	3	3	3	93.67%	94.94%	92.41%	96.20%	96.20%	96.20%
Average Error										93.92%	93.67%	92.91%	94.18%	94.94%	95.44%

Table A 19. Feature Selection with Mutual Information on Breast Cancer Data

FS: Mutual Information w/ Equal Width on breast cancer Data (w=10)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	28	11	22	7	5	8	4	4	4	91.14%	93.67%	89.87%	94.94%	94.94%	94.94%
Fold2	28	11	22	7	9	6	7	7	7	91.14%	88.61%	92.41%	91.14%	91.14%	91.14%
Fold3	28	11	22	8	8	7	6	6	6	89.87%	89.87%	91.14%	92.41%	92.41%	92.41%
Fold4	23	2	28	6	2	4	3	3	3	92.41%	97.47%	94.94%	96.20%	96.20%	96.20%
Fold5	23	22	28	6	4	4	3	3	3	92.41%	94.94%	94.94%	96.20%	96.20%	96.20%
Average Error										91.39%	92.91%	92.66%	94.18%	94.18%	94.18%

Table A 20. Feature Selection with Mutual Information on Breast Cancer Data

FS: Mutual Information w/ Equal Width on breast cancer Data (w=15)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	23	27	14	9	7	10	7	8	8	88.61%	91.14%	87.34%	91.14%	89.87%	89.87%
Fold2	23	27	14	9	7	9	10	10	10	88.61%	91.14%	88.61%	87.34%	87.34%	87.34%
Fold3	23	7	14	13	7	7	7	7	7	83.54%	91.14%	91.14%	91.14%	91.14%	91.14%
Fold4	23	28	14	6	5	6	3	3	3	92.41%	93.67%	92.41%	96.20%	96.20%	96.20%
Fold5	23	27	14	11	4	8	8	8	8	86.08%	94.94%	89.87%	89.87%	89.87%	89.87%
Average Error										87.85%	92.41%	89.87%	91.14%	90.89%	90.89%

Table A 21. Feature Selection with Mutual Information on Breast Cancer Data

FS: Mutual Information w/ Equal Width on breast cancer Data (w=20)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	23	27	14	9	7	10	7	8	8	88.61%	91.14%	87.34%	91.14%	89.87%	89.87%
Fold2	23	27	14	9	7	9	10	10	10	88.61%	91.14%	88.61%	87.34%	87.34%	87.34%
Fold3	23	27	14	10	5	7	6	5	5	87.34%	93.67%	91.14%	92.41%	93.67%	93.67%
Fold4	23	8	14	5	6	6	4	4	4	93.67%	92.41%	92.41%	94.94%	94.94%	94.94%
Fold5	23	27	14	11	4	8	8	8	8	86.08%	94.94%	89.87%	89.87%	89.87%	89.87%
Average Error										88.86%	92.66%	89.87%	91.14%	91.14%	91.14%

Table A 22. Feature Selection with Mutual Information on Breast Cancer Data

FS: Mutual Information w/ Equal Frequency on breast cancer Data (f=5)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	23	22	25	9	5	6	3	3	3	88.61%	93.67%	92.41%	96.20%	96.20%	96.20%
Fold2	23	25	2	7	8	11	7	7	7	91.14%	89.87%	86.08%	91.14%	91.14%	91.14%
Fold3	8	22	21	6	5	7	4	4	4	92.41%	93.67%	91.14%	94.94%	94.94%	94.94%
Fold4	23	25	2	3	3	5	2	1	1	96.20%	96.20%	93.67%	97.47%	98.73%	98.73%
Fold5	23	25	2	4	4	5	2	2	2	94.94%	94.94%	93.67%	97.47%	97.47%	97.47%
Average Error										92.66%	93.67%	91.39%	95.44%	95.70%	95.70%

Table A 23. Feature Selection with Mutual Information on Breast Cancer Data

FS: Mutual Information w/ Equal Frequency on breast cancer Data (f=10)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	23	22	28	8	5	6	3	3	3	89.87%	93.67%	92.41%	96.20%	96.20%	96.20%
Fold2	23	22	8	7	8	7	5	5	5	91.14%	89.87%	91.14%	93.67%	93.67%	93.67%
Fold3	23	22	8	4	4	7	4	4	4	94.94%	94.94%	91.14%	94.94%	94.94%	94.94%
Fold4	23	28	2	6	2	4	3	3	3	92.41%	97.47%	94.94%	96.20%	96.20%	96.20%
Fold5	23	27	14	11	4	8	8	8	8	86.08%	94.94%	89.87%	89.87%	89.87%	89.87%
Average Error										90.89%	94.18%	91.90%	94.18%	94.18%	94.18%

Table A 24. Feature Selection with Mutual Information on Breast Cancer Data

FS: Mutual Information w/ Equal Frequency on breast cancer Data (f=15)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	23	27	14	9	7	10	7	8	8	88.61%	91.14%	87.34%	91.14%	89.87%	89.87%
Fold2	23	28	14	17	7	8	6	4	4	78.48%	91.14%	89.87%	92.41%	94.94%	94.94%
Fold3	23	28	14	7	5	5	6	7	6	91.14%	93.67%	93.67%	92.41%	91.14%	92.41%
Fold4	23	28	14	6	5	6	3	3	3	92.41%	93.67%	92.41%	96.20%	96.20%	96.20%
Fold5	23	27	8	7	4	7	9	7	7	91.14%	94.94%	91.14%	88.61%	91.14%	91.14%
Average Error										88.35%	92.91%	90.89%	92.15%	92.66%	92.91%

Table A 25. Feature Selection with Mutual Information on Breast Cancer Data

FS: Mutual Information w/ Equal Frequency on breast cancer Data (f=20)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	23	28	14	12	5	6	5	5	5	84.81%	93.67%	92.41%	93.67%	93.67%	93.67%
Fold2	23	28	14	17	7	8	6	4	4	78.48%	91.14%	89.87%	92.41%	94.94%	94.94%
Fold3	23	8	14	19	7	7	8	8	8	75.95%	91.14%	91.14%	89.87%	89.87%	89.87%
Fold4	23	28	14	6	5	6	3	3	3	92.41%	93.67%	92.41%	96.20%	96.20%	96.20%
Fold5	23	28	14	11	5	4	7	8	8	86.08%	93.67%	94.94%	91.14%	89.87%	89.87%
Average Error										83.54%	92.66%	92.15%	92.66%	92.91%	92.91%

Table A 26. Feature Selection with Stepwise on Breast Cancer Data

FS: Stepwise on breast cancer Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	21	28	22	7	5	4	2	3	3	93.81%	95.58%	96.46%	98.23%	97.35%	97.35%
Fold2	21	8	26	9	9	8	7	7	7	92.04%	92.04%	92.92%	93.81%	93.81%	93.81%
Fold3	21	27	11	8	3	3	6	5	4	92.92%	97.35%	97.35%	94.69%	95.58%	96.46%
Fold4	21	27	22	4	5	5	3	2	3	96.46%	95.58%	95.58%	97.35%	98.23%	97.35%
Fold5	21	30	29	5	8	9	4	5	5	95.58%	92.92%	92.04%	96.46%	95.58%	95.58%
Average Error										94.16%	94.69%	94.87%	96.11%	96.11%	96.11%

Table A 27. Feature Selection with Welch t-Statistics on Breast Cancer Data

FS: Welch t-Statistics I on Breast Cancer Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	28	23	21	8	6	5	6	5	5	92.92%	94.69%	95.58%	94.69%	95.58%	95.58%
Fold2	28	23	21	14	17	6	7	7	6	87.61%	84.96%	94.69%	93.81%	93.81%	94.69%
Fold3	28	8	23	12	6	7	8	8	7	89.38%	94.69%	93.81%	92.92%	92.92%	93.81%
Fold4	28	23	21	4	6	6	5	5	5	96.46%	94.69%	94.69%	95.58%	95.58%	95.58%
Fold5	28	23	21	10	5	5	7	7	7	91.15%	95.58%	95.58%	93.81%	93.81%	93.81%
Average Error										91.50%	92.92%	94.87%	94.16%	94.34%	94.69%

Table A 28. Feature Selection with Fisher Correlation Score on Breast Cancer Data

FS: Fisher Correlation Score I on Breast Cancer Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	28	23	21	8	6	5	6	5	5	92.92%	94.69%	95.58%	94.69%	95.58%	95.58%
Fold2	28	23	21	14	17	6	7	7	6	87.61%	84.96%	94.69%	93.81%	93.81%	94.69%
Fold3	28	8	23	12	6	7	8	8	7	89.38%	94.69%	93.81%	92.92%	92.92%	93.81%
Fold4	28	23	21	4	6	6	5	5	5	96.46%	94.69%	94.69%	95.58%	95.58%	95.58%
Fold5	28	23	21	10	5	5	7	7	7	91.15%	95.58%	95.58%	93.81%	93.81%	93.81%
Average Error										91.50%	92.92%	94.87%	94.16%	94.34%	94.69%

Table A 29. Feature Selection with Independently Consistent Expression on Breast Cancer Data

FS: Independently consistent Expression on Breast Cancer Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	14	24	4	10	10	9	10	8	7	91.15%	91.15%	92.04%	91.15%	92.92%	93.81%
Fold2	14	24	4	16	11	9	10	10	9	85.84%	90.27%	92.04%	91.15%	91.15%	92.04%
Fold3	14	24	4	12	11	9	7	7	6	89.38%	90.27%	92.04%	93.81%	93.81%	94.69%
Fold4	14	24	4	13	11	9	6	6	6	88.50%	90.27%	92.04%	94.69%	94.69%	94.69%
Fold5	14	24	4	12	14	7	7	7	7	89.38%	87.61%	93.81%	93.81%	93.81%	93.81%
Average Error										88.85%	89.91%	92.39%	92.92%	93.27%	93.81%

Table A 30. Feature Selection with Mean Difference Score on Breast Cancer Data

FS: Mean Difference Score on Breast Cancer Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	24	14	23	12	8	8	11	11	10	89.38%	92.92%	92.92%	90.27%	90.27%	91.15%
Fold2	14	24	8	8	8	7	9	8	8	92.92%	92.92%	93.81%	92.04%	92.92%	92.92%
Fold3	14	24	8	5	6	7	10	9	10	95.58%	94.69%	93.81%	91.15%	92.04%	91.15%
Fold4	24	14	23	13	8	8	7	6	6	88.50%	92.92%	92.92%	93.81%	94.69%	94.69%
Fold5	24	14	23	19	12	10	9	8	9	83.19%	89.38%	91.15%	92.04%	92.92%	92.04%
Average Error										89.91%	92.57%	92.92%	91.86%	92.57%	92.39%

Table A 31. Feature Selection with Average Difference Score on Breast Cancer Data

FS: Average Difference Score on Breast Cancer Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	3	6	1	9	9	9	14	13	13	92.04%	92.04%	92.04%	87.61%	88.50%	88.50%
Fold2	3	24	1	16	10	6	10	9	9	85.84%	91.15%	94.69%	91.15%	92.04%	92.04%
Fold3	24	23	3	5	7	4	3	3	4	95.58%	93.81%	96.46%	97.35%	97.35%	96.46%
Fold4	14	13	3	15	12	12	12	13	12	86.73%	89.38%	89.38%	89.38%	88.50%	89.38%
Fold5	3	1	2	37	11	11	11	11	11	67.26%	90.27%	90.27%	90.27%	90.27%	90.27%
Average Error										85.49%	91.33%	92.57%	91.15%	91.33%	91.33%

Table A 32. Feature Selection with Relief Algorithm on Breast Cancer Data

FS: Relief Algorithm on Breast Cancer Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	21	1	3	11	10	8	8	9	8	90.27%	91.15%	92.92%	92.92%	92.04%	92.92%
Fold2	21	23	3	7	9	8	8	8	8	93.81%	92.04%	92.92%	92.92%	92.92%	92.92%
Fold3	3	1	21	12	10	4	7	5	4	89.38%	91.15%	96.46%	93.81%	95.58%	96.46%
Fold4	1	3	21	8	11	7	7	5		92.92%	90.27%	93.81%	93.81%	95.58%	100.00%
Fold5	3	1	21	10	15	9	6	7	7	91.15%	86.73%	92.04%	94.69%	93.81%	93.81%
Average Error										91.50%	90.27%	93.63%	93.63%	93.98%	95.22%

B. Results on PIMA Diabetes Dataset

Table B 1. Average errors of five folds on PIMA Diabetes data from EQ. 4.9

i	j	k	Average Error Bases on Equation (4.9)	Standard Deviation
2	6	8	37.4	2.881
2	3	6	37.6	4.6152
1	2	8	38.4	3.9749
2	5	6	38.4	6.8044
1	2	6	39	5.099
2	6	7	39.6	3.3615
2	4	8	42.2	8.0125
2	5	7	43.2	5.8052
2	3	4	43.2	6.7231
1	2	7	43.6	8.0187
2	3	8	44	10.2956
2	7	8	44.6	4.8785
2	4	6	45.2	6.4187
1	2	3	45.6	7.0922
2	3	5	45.6	9.5289
2	5	8	46.2	7.5631
1	2	4	46.4	2.9665
2	3	7	46.4	9.0443
1	6	7	46.6	7.0214
4	7	8	49.6	4.5056
6	7	8	49.6	4.7223
1	3	6	50	2.4495
4	6	7	50.8	6.3008
2	4	7	51	8.124
5	6	7	51.4	4.0373
2	4	5	51.6	5.9414
1	6	8	51.8	7.3959
5	6	8	52	6.364
3	7	8	52.2	2.6833
4	6	8	52.2	4.7645
1	7	8	52.8	6.0992
1	2	5	53	14.089
3	6	7	53.6	8.4735
4	5	6	54	4.1833
3	5	6	54.4	7.9246
3	4	8	54.6	4.7223
3	5	8	54.8	6.0581
1	5	6	55	7.0711
3	6	8	55.2	4.6583
5	7	8	55.4	5.5498
1	5	7	55.4	7.2319
1	3	8	55.6	7.7006
3	5	7	55.8	6.6106
1	4	6	57	9.1924
3	4	7	57.6	4.827
4	5	8	58	5.6125
1	4	7	58	8.6891
1	3	7	58.2	6.3403
4	5	7	59.4	5.9414
1	3	4	61	7.3144

Table B 2. Average errors of five folds on Breast Cancer data from EQ. 4.10

i	j	k	Average Error Bases on Equation (4.10)	Standard Deviation
1	2	6	34	4.4159
2	6	7	34.6	4.3359
2	7	8	36	1.5811
2	5	8	36.4	1.6733
1	2	5	36.6	2.881
1	2	8	36.8	2.3875
2	3	8	36.8	2.9496
2	4	6	37.2	3.7683
1	2	3	37.4	2.6077
2	6	8	37.6	2.6077
2	3	6	37.6	2.7019
2	4	7	37.8	1.6432
2	4	8	38	4.1231
1	2	4	38.2	1.0954
2	3	5	38.2	1.9235
1	2	7	39	2.5495
2	3	7	39.6	2.3022
2	5	6	39.8	3.1145
2	4	5	40	6.8191
2	5	7	40.2	4.2071
2	3	4	41	3.5355
3	6	8	43.6	2.3022
6	7	8	44.4	4.5056
1	6	8	44.6	3.5071
1	6	7	44.6	5.1769
4	6	7	44.8	1.3038
5	6	7	45.2	3.5637
3	6	7	45.2	3.6332
4	6	8	45.4	5.4129
5	6	8	46	3.1623
5	7	8	46.2	3.5637
1	5	6	46.2	4.9193
4	7	8	46.4	2.7928
3	5	6	47	2.8284
1	5	8	47.2	2.1679
3	5	8	47.6	3.0496
1	3	6	47.6	3.2094
4	5	6	48	2.4495
3	4	7	48	3
3	7	8	48.2	2.5884
4	5	8	48.2	3.4205
1	4	7	48.8	2.7749
3	4	8	48.8	4.0866
1	3	5	49.2	2.3875
1	5	7	49.2	3.7014
1	4	5	49.4	3.5071
1	7	8	49.4	4.6152
1	3	7	49.8	3.3466
3	5	7	49.8	3.7683
1	3	8	49.8	4.7645

Table B 3. Average errors of five folds on PIMA Diabetes data from ANFIS

i	j	k	Average Error based on ANFIS	Standard Deviation
2	7	8	34	4
2	6	8	34.2	2.5884
2	3	8	35	3.5355
2	4	8	35.2	4.3243
1	2	8	35.4	2.51
2	6	7	36.2	3.1145
2	5	8	36.4	4.219
1	2	3	36.4	4.2778
2	3	7	36.6	2.51
1	2	6	36.6	3.5071
1	2	4	37.2	2.3875
2	3	6	37.4	1.1402
1	2	7	37.4	1.3416
2	4	6	37.4	1.6733
2	4	5	37.6	3.0496
2	5	6	38	2
1	2	5	38	2.3452
2	4	7	38.4	2.51
2	3	4	38.6	2.51
2	5	7	38.6	2.7019
2	3	5	39.2	1.7889
5	7	8	43.2	0.8367
6	7	8	43.2	3.1937
3	5	8	43.8	2.2804
5	6	8	43.8	2.49
4	5	8	43.8	2.9496
1	5	8	44.8	2.9496
1	6	7	45	2.1213
1	6	8	45.6	2.7019
3	6	8	45.6	3.2094
4	7	8	46	1.2247
4	6	8	46	1.5811
1	5	7	46.4	3.9115
1	4	6	46.6	1.9494
1	3	5	46.6	3.2094
4	5	6	46.8	3.3466
3	5	6	47	2.2361
4	6	7	47	3.5355
5	6	7	47	4.6368
1	5	6	47.4	3.5071
1	3	8	47.6	1.8166
3	7	8	47.6	1.9494
3	4	8	47.6	3.9115
3	4	5	48	3.6742
4	5	7	48	4.4159
1	7	8	48.2	1.9235
1	4	7	48.2	2.1679
3	6	7	48.2	3.7014
3	5	7	48.4	3.9115
1	4	5	48.8	2.7749

Table B 4. Average errors of five folds on PIMA Diabetes data from Fuzzy SNN

i	j	k	Average Error based on Fuzzy KNN (N=5)	Standard Deviation
2	6	7	37.2	2.1679
1	2	8	37.6	3.2094
2	6	8	38	4.8477
2	4	7	38.6	3.1305
2	5	8	38.8	3.9623
2	7	8	39.4	3.4351
1	2	7	40.4	2.4083
2	3	8	40.6	2.51
1	2	6	40.6	3.2094
1	2	5	41.4	5.1284
2	3	5	41.6	1.3416
1	2	4	41.6	6.3482
2	3	6	41.8	2.6833
2	4	6	42	2.9155
2	5	7	42	3.3912
2	5	6	42.4	3.4351
6	7	8	43.2	1.4832
1	2	3	43.2	2.9496
2	4	8	43.2	6.0992
2	4	5	44	3.4641
5	6	8	44	4
2	3	4	44.4	4.3359
2	3	7	45.2	2.7749
4	5	8	46	4.3589
1	3	5	46.8	1.4832
1	5	8	46.8	3.8987
1	6	8	47	3.873
1	6	7	47.2	3.5637
4	6	7	47.8	2.8636
4	5	6	47.8	3.1145
5	6	7	47.8	3.2711
5	7	8	47.8	6.9426
1	5	7	49.2	3.5637
3	6	8	49.6	2.6077
4	6	8	50.6	2.4083
1	3	6	50.6	3.2094
4	7	8	50.6	4.3359
3	4	5	50.8	2.49
1	5	6	50.8	4.8166
4	5	7	51	1.8708
3	5	8	51.2	2.2804
3	6	7	51.4	1.3416
1	7	8	51.6	4.8785
3	5	6	52	4.8477
3	7	8	52.4	2.881
1	4	5	52.6	1.6733
3	4	6	52.6	3.5777
1	3	4	52.8	2.8636
3	4	8	53	2.5495
1	3	7	53.4	3.5071

Table B 5. Average errors of five folds on PIMA Diabetes data from Fuzzy 10NN

i	j	k	Average Error based on Fuzzy KNN (N=10)	Standard Deviation
2	6	8	35.6	4.219
2	4	7	36.4	3.9115
2	6	7	36.6	2.881
2	7	8	36.8	2.2804
1	2	8	37.2	3.4205
1	2	6	38.6	1.1402
2	3	8	38.8	2.9496
2	3	6	39	2.3452
2	4	6	39	2.5495
2	5	8	39	4
1	2	5	39	4.062
1	2	7	39.6	2.0736
2	5	7	39.8	3.1145
2	5	6	40	1.8708
2	3	7	40.2	3.0332
2	3	5	40.8	2.2804
1	2	4	41.4	6.0663
2	3	4	41.6	2.7019
1	2	3	41.8	4.5497
2	4	8	42.6	5.1769
6	7	8	42.8	2.1679
5	6	8	43	3.5355
2	4	5	43	3.937
4	5	8	44.2	4.8166
1	5	8	45.6	4.0988
5	6	7	46.2	3.1145
1	3	5	46.2	3.3466
1	6	8	46.6	5.8992
4	5	6	46.8	2.7749
3	6	8	46.8	4.0866
5	7	8	46.8	6.0166
1	6	7	47	2.8284
4	6	7	47.6	2.0736
1	5	7	47.8	2.2804
4	7	8	48.4	2.9665
3	6	7	48.6	2.7019
4	6	8	49.4	2.881
1	5	6	49.6	4.1593
1	3	6	49.6	4.3359
4	5	7	50	1.2247
3	4	7	50.4	3.6469
3	4	5	50.6	2.4083
3	5	8	50.8	1.0954
1	7	8	50.8	5.0695
1	4	7	51	3.8079
3	4	8	52	2.1213
3	7	8	52.2	3.3466
1	4	5	52.4	0.5477
3	5	7	52.4	3.5777
1	3	4	52.6	3.2863

Table B 6. Average errors of five folds on PIMA Diabetes data from Fuzzy 15NN

i	j	k	Average Error based on Fuzzy KNN (N=15)	Standard Deviation
2	4	7	35.4	3.9749
2	6	8	35.8	3.5637
2	7	8	36	2.5495
2	6	7	36.4	3.5777
1	2	8	36.8	3.7683
2	3	6	37.2	2.5884
1	2	6	38	1.2247
2	3	7	38.2	3.0332
2	4	6	38.4	2.7019
2	5	8	38.4	3.9115
1	2	7	38.6	1.5166
2	3	8	38.6	4.3932
1	2	5	39	4.1833
2	5	7	39.2	3.6332
2	5	6	39.8	1.7889
2	3	4	40.2	2.1679
2	3	5	40.2	2.3875
1	2	4	40.2	6.3008
1	2	3	41	3.6742
2	4	8	41.2	5.933
6	7	8	42	2.9155
2	4	5	42.2	2.7749
5	6	8	42.4	3.2094
4	5	8	44.4	5.5498
1	6	7	45.6	2.3022
1	3	5	45.8	2.5884
4	5	6	45.8	3.0332
1	5	8	45.8	3.7014
5	7	8	45.8	5.8907
5	6	7	46.4	3.5777
1	6	8	46.4	3.7815
3	6	8	46.8	3.6332
4	6	7	47.4	2.7019
4	7	8	47.8	2.0494
1	5	7	47.8	2.9496
4	6	8	48.2	3.4205
3	6	7	48.6	2.1909
1	5	6	49	3.937
4	5	7	49.2	2.0494
1	3	6	49.2	5.0695
3	4	5	49.8	2.8636
3	4	7	50	3.873
3	5	8	50.4	1.9494
1	4	5	51.2	1.0954
1	3	4	51.2	3.5637
1	7	8	51.2	4.7645
3	5	7	51.6	4.0373
3	7	8	52	2.9155
1	3	8	52	3.1623
1	4	7	52	4.3589

Table B 7. Feature Selection with Mutual Correlation on PIMA Diabetes Data

FS: mutual correlation (Tsai and Chiu) PIMA Diabetes Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	6	8	33	36	35	31	33	32	78.43%	76.47%	77.12%	79.74%	78.43%	79.08%
Fold2	2	6	1	45	34	41	44	39	39	70.59%	77.78%	73.20%	71.24%	74.51%	74.51%
Fold3	2	6	8	37	36	38	43	41	40	75.82%	76.47%	75.16%	71.90%	73.20%	73.86%
Fold4	2	6	1	33	34	36	37	37	36	78.43%	77.78%	76.47%	75.82%	75.82%	76.47%
Fold5	2	6	1	43	39	32	44	40	39	71.90%	74.51%	79.08%	71.24%	73.86%	74.51%
Average Error										75.03%	76.60%	76.21%	73.99%	75.16%	75.69%

Table B 8. Feature Selection with Mutual Correlation on PIMA Diabetes Data

FS: mutual correlation (Haindl et al) PIMA Diabetes Data (Liao)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	1	5	7	55	50	41	53	50	51	64.05%	67.32%	73.20%	65.36%	67.32%	66.67%
Fold2	1	3	7	52	51	52	55	55	54	66.01%	66.67%	66.01%	64.05%	64.05%	64.71%
Fold3	4	7	8	44	44	45	49	47	48	71.24%	71.24%	70.59%	67.97%	69.28%	68.63%
Fold4	1	4	7	50	50	51	54	52	52	67.32%	67.32%	66.67%	64.71%	66.01%	66.01%
Fold5	1	4	7	51	45	46	52	47	47	66.67%	70.59%	69.93%	66.01%	69.28%	69.28%
Average Error										67.06%	68.63%	69.28%	65.62%	67.19%	67.06%

Table B 9. Feature Selection with Mutual Correlation on PIMA Diabetes Data

FS: mutual correlation (Park et al) PIMA Diabetes Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	6	8	33	36	35	31	33	32	78.43%	76.47%	77.12%	79.74%	78.43%	79.08%
Fold2	1	2	6	45	34	41	44	39	39	70.59%	77.78%	73.20%	71.24%	74.51%	74.51%
Fold3	2	6	8	37	36	38	43	41	40	75.82%	76.47%	75.16%	71.90%	73.20%	73.86%
Fold4	2	6	8	38	42	31	38	34	35	75.16%	72.55%	79.74%	75.16%	77.78%	77.12%
Fold5	1	2	6	43	39	32	44	40	39	71.90%	74.51%	79.08%	71.24%	73.86%	74.51%
Average Error										74.38%	75.56%	76.86%	73.86%	75.56%	75.82%

Table B 10. Feature Selection with Mutual Information on PIMA Diabetes Data

FS: Mutual Information w/ Equal Width on PIMA diabetes Data (w=5)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	1	6	35	27	35	39	38	38	55.70%	65.82%	55.70%	50.63%	51.90%	51.90%
Fold2	2	6	7	42	36	34	40	41	41	46.84%	54.43%	56.96%	49.37%	48.10%	48.10%
Fold3	2	6	1	39	36	39	39	39	38	50.63%	54.43%	50.63%	50.63%	50.63%	51.90%
Fold4	2	6	8	38	42	31	38	34	35	51.90%	46.84%	60.76%	51.90%	56.96%	55.70%
Fold5	2	6	1	43	39	32	44	40	39	45.57%	50.63%	59.49%	44.30%	49.37%	50.63%
Average Error										50.13%	54.43%	56.71%	49.37%	51.39%	51.65%

Table B 11. Feature Selection with Mutual Information on PIMA Diabetes Data

FS: Mutual Information w/ Equal Width on PIMA diabetes Data (w=10)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	6	8	33	36	35	31	33	32	58.23%	54.43%	55.70%	60.76%	58.23%	59.49%
Fold2	2	6	8	41	36	34	36	31	33	48.10%	54.43%	56.96%	54.43%	60.76%	58.23%
Fold3	2	6	8	37	36	38	43	41	40	53.16%	54.43%	51.90%	45.57%	48.10%	49.37%
Fold4	2	6	8	38	42	31	38	34	35	51.90%	46.84%	60.76%	51.90%	56.96%	55.70%
Fold5	2	6	8	38	38	33	42	39	39	51.90%	51.90%	58.23%	46.84%	50.63%	50.63%
Average Error										52.66%	52.41%	56.71%	51.90%	54.94%	54.68%

Table B 12. Feature Selection with Mutual Information on PIMA Diabetes Data

FS: Mutual Information w/ Equal Width on PIMA diabetes Data (w=15)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	4	7	55	35	39	38	37	36	30.38%	55.70%	50.63%	51.90%	53.16%	54.43%
Fold2	2	6	7	42	36	34	40	41	41	46.84%	54.43%	56.96%	49.37%	48.10%	48.10%
Fold3	2	6	7	39	38	40	38	37	39	50.63%	51.90%	49.37%	51.90%	53.16%	50.63%
Fold4	2	6	7	37	36	39	37	34	35	53.16%	54.43%	50.63%	53.16%	56.96%	55.70%
Fold5	2	6	8	38	38	33	42	39	39	51.90%	51.90%	58.23%	46.84%	50.63%	50.63%
Average Error										46.58%	53.67%	53.16%	50.63%	52.41%	51.90%

Table B 13. Feature Selection with Mutual Information on PIMA Diabetes Data

FS: Mutual Information w/ Equal Width on PIMA diabetes Data (w=20)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	4	7	55	35	39	38	37	36	30.38%	55.70%	50.63%	51.90%	53.16%	54.43%
Fold2	2	6	7	42	36	34	40	41	41	46.84%	54.43%	56.96%	49.37%	48.10%	48.10%
Fold3	2	4	1	49	40	39	49	48	48	37.97%	49.37%	50.63%	37.97%	39.24%	39.24%
Fold4	2	6	7	37	36	39	37	34	35	53.16%	54.43%	50.63%	53.16%	56.96%	55.70%
Fold5	2	6	5	41	39	39	41	41	40	48.10%	50.63%	50.63%	48.10%	48.10%	49.37%
Average Error										43.29%	52.91%	51.90%	48.10%	49.11%	49.37%

Table B 14. Feature Selection with Mutual Information on PIMA Diabetes Data

FS: Mutual Information w/ Equal Frequency on PIMA diabetes Data (f=5)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	5	8	2	53	36	38	38	40	39	32.91%	54.43%	51.90%	51.90%	49.37%	50.63%
Fold2	5	1	2	70	32	40	43	39	40	11.39%	59.49%	49.37%	45.57%	50.63%	49.37%
Fold3	5	1	2	59	39	36	46	43	41	25.32%	50.63%	54.43%	41.77%	45.57%	48.10%
Fold4	5	1	2	59	37	35	46	43	44	25.32%	53.16%	55.70%	41.77%	45.57%	44.30%
Fold5	5	1	2	41	39	39	35	34	33	48.10%	50.63%	50.63%	55.70%	56.96%	58.23%
Average Error										28.61%	53.67%	52.41%	47.34%	49.62%	50.13%

Table B 15. Feature Selection with Mutual Information on PIMA Diabetes Data

FS: Mutual Information w/ Equal Frequency on PIMA diabetes Data (f=10)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	5	8	7	52	46	44	54	53	50	34.18%	41.77%	44.30%	31.65%	32.91%	36.71%
Fold2	5	1	2	70	32	40	43	39	40	11.39%	59.49%	49.37%	45.57%	50.63%	49.37%
Fold3	5	1	2	59	39	36	46	43	41	25.32%	50.63%	54.43%	41.77%	45.57%	48.10%
Fold4	5	8	7	62	46	44	42	41	39	21.52%	41.77%	44.30%	46.84%	48.10%	50.63%
Fold5	5	6	2	41	39	39	41	41	40	48.10%	50.63%	50.63%	48.10%	48.10%	49.37%
Average Error										28.10%	48.86%	48.61%	42.78%	45.06%	46.84%

Table B 16. Feature Selection with Mutual Information on PIMA Diabetes Data

FS: Mutual Information w/ Equal Frequency on PIMA diabetes Data (f=15)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	5	8	2	53	36	38	38	40	39	32.91%	54.43%	51.90%	51.90%	49.37%	50.63%
Fold2	5	1	2	70	32	40	43	39	40	11.39%	59.49%	49.37%	45.57%	50.63%	49.37%
Fold3	5	6	8	59	45	46	43	42	42	25.32%	43.04%	41.77%	45.57%	46.84%	46.84%
Fold4	5	6	2	41	45	35	37	38	37	48.10%	43.04%	55.70%	53.16%	51.90%	53.16%
Fold5	5	7	1	58	43	46	44	44	43	26.58%	45.57%	41.77%	44.30%	44.30%	45.57%
Average Error										28.86%	49.11%	48.10%	48.10%	48.61%	49.11%

Table B 17. Feature Selection with Mutual Information on PIMA Diabetes Data

FS: Mutual Information w/ Equal Frequency on PIMA diabetes Data (f=20)(mid)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	5	1	2	36	36	40	37	36	37	54.43%	54.43%	49.37%	53.16%	54.43%	53.16%
Fold2	5	1	7	47	50	52	49	48	49	40.51%	36.71%	34.18%	37.97%	39.24%	37.97%
Fold3	5	1	2	59	39	36	46	43	41	25.32%	50.63%	54.43%	41.77%	45.57%	48.10%
Fold4	5	1	7	66	50	47	48	48	48	16.46%	36.71%	40.51%	39.24%	39.24%	39.24%
Fold5	5	1	2	41	39	39	35	34	33	48.10%	50.63%	50.63%	55.70%	56.96%	58.23%
Average Error										36.96%	45.82%	45.82%	45.57%	47.09%	47.34%

Table B 18. Feature Selection with Mutual Information on PIMA Diabetes Data

FS: Mutual Information w/ Equal Width on PIMA diabetes Data (w=5)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	1	6	35	27	35	39	38	38	55.70%	65.82%	55.70%	50.63%	51.90%	51.90%
Fold2	2	6	7	53	49	47	48	50	48	32.91%	37.97%	40.51%	39.24%	36.71%	39.24%
Fold3	2	6	1	39	36	39	39	39	38	50.63%	54.43%	50.63%	50.63%	50.63%	51.90%
Fold4	2	6	8	38	42	31	38	34	35	51.90%	46.84%	60.76%	51.90%	56.96%	55.70%
Fold5	2	6	1	43	39	32	44	40	39	45.57%	50.63%	59.49%	44.30%	49.37%	50.63%
Average Error										47.34%	51.14%	53.42%	47.34%	49.11%	49.87%

Table B 19. Feature Selection with Mutual Information on PIMA Diabetes Data

FS: Mutual Information w/ Equal Width on PIMA diabetes Data (w=10)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	6	8	33	36	35	31	33	32	58.23%	54.43%	55.70%	60.76%	58.23%	59.49%
Fold2	2	6	8	41	36	34	36	31	33	48.10%	54.43%	56.96%	54.43%	60.76%	58.23%
Fold3	2	6	8	37	36	38	43	41	40	53.16%	54.43%	51.90%	45.57%	48.10%	49.37%
Fold4	2	6	8	38	42	31	38	34	35	51.90%	46.84%	60.76%	51.90%	56.96%	55.70%
Fold5	2	6	8	38	38	33	42	39	39	51.90%	51.90%	58.23%	46.84%	50.63%	50.63%
Average Error										52.66%	52.41%	56.71%	51.90%	54.94%	54.68%

Table B 20. Feature Selection with Mutual Information on PIMA Diabetes Data

FS: Mutual Information w/ Equal Width on PIMA diabetes Data (w=15)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	6	8	33	36	35	31	33	32	58.23%	54.43%	55.70%	60.76%	58.23%	59.49%
Fold2	2	6	8	41	36	34	36	31	33	48.10%	54.43%	56.96%	54.43%	60.76%	58.23%
Fold3	2	6	8	37	36	38	43	41	40	53.16%	54.43%	51.90%	45.57%	48.10%	49.37%
Fold4	2	6	8	38	42	31	38	34	35	51.90%	46.84%	60.76%	51.90%	56.96%	55.70%
Fold5	2	6	8	38	38	33	42	39	39	51.90%	51.90%	58.23%	46.84%	50.63%	50.63%
Average Error										52.66%	52.41%	56.71%	51.90%	54.94%	54.68%

Table B 21. Feature Selection with Mutual Information on PIMA Diabetes Data

FS: Mutual Information w/ Equal Width on PIMA diabetes Data (w=20)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	6	8	33	36	35	31	33	32	58.23%	54.43%	55.70%	60.76%	58.23%	59.49%
Fold2	2	6	8	41	36	34	36	31	33	48.10%	54.43%	56.96%	54.43%	60.76%	58.23%
Fold3	2	6	8	37	36	38	43	41	40	53.16%	54.43%	51.90%	45.57%	48.10%	49.37%
Fold4	2	6	8	38	42	31	38	34	35	51.90%	46.84%	60.76%	51.90%	56.96%	55.70%
Fold5	2	6	8	38	38	33	42	39	39	51.90%	51.90%	58.23%	46.84%	50.63%	50.63%
Average Error										52.66%	52.41%	56.71%	51.90%	54.94%	54.68%

Table B 22. Feature Selection with Mutual Information on PIMA Diabetes Data

FS: Mutual Information w/ Equal Frequency on PIMA diabetes Data (f=5)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	5	8	2	53	36	38	38	40	39	32.91%	54.43%	51.90%	51.90%	49.37%	50.63%
Fold2	5	1	6	53	50	51	56	56	54	32.91%	36.71%	35.44%	29.11%	29.11%	31.65%
Fold3	5	1	2	59	39	36	46	43	41	25.32%	50.63%	54.43%	41.77%	45.57%	48.10%
Fold4	5	1	2	59	37	35	46	43	44	25.32%	53.16%	55.70%	41.77%	45.57%	44.30%
Fold5	5	6	2	41	39	39	41	41	40	48.10%	50.63%	50.63%	48.10%	48.10%	49.37%
Average Error										32.91%	49.11%	49.62%	42.53%	43.54%	44.81%

Table B 23. Feature Selection with Mutual Information on PIMA Diabetes Data

FS: Mutual Information w/ Equal Frequency on PIMA diabetes Data (f=10)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	5	1	2	36	36	40	37	36	37	54.43%	54.43%	49.37%	53.16%	54.43%	53.16%
Fold2	5	1	2	70	32	40	43	39	40	11.39%	59.49%	49.37%	45.57%	50.63%	49.37%
Fold3	5	2	1	59	39	36	46	43	41	25.32%	50.63%	54.43%	41.77%	45.57%	48.10%
Fold4	5	2	6	41	45	35	37	38	37	48.10%	43.04%	55.70%	53.16%	51.90%	53.16%
Fold5	5	2	6	41	39	39	41	41	40	48.10%	50.63%	50.63%	48.10%	48.10%	49.37%
Average Error										37.47%	51.65%	51.90%	48.35%	50.13%	50.63%

Table B 24. Feature Selection with Mutual Information on PIMA Diabetes Data

FS: Mutual Information w/ Equal Frequency on PIMA diabetes Data (f=15)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	5	1	2	36	36	40	37	36	37	54.43%	54.43%	49.37%	53.16%	54.43%	53.16%
Fold2	5	3	1	77	48	50	46	41	43	2.53%	39.24%	36.71%	41.77%	48.10%	45.57%
Fold3	5	6	8	59	45	46	43	42	42	25.32%	43.04%	41.77%	45.57%	46.84%	46.84%
Fold4	5	2	1	59	37	35	46	43	44	25.32%	53.16%	55.70%	41.77%	45.57%	44.30%
Fold5	5	2	1	41	39	39	35	34	33	48.10%	50.63%	50.63%	55.70%	56.96%	58.23%
Average Error										31.14%	48.10%	46.84%	47.59%	50.38%	49.62%

Table B 25. Feature Selection with Mutual Information on PIMA Diabetes Data

FS: Mutual Information w/ Equal Frequency on PIMA diabetes Data (f=20)(miq)				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	5	1	2	36	36	40	37	36	37	54.43%	54.43%	49.37%	53.16%	54.43%	53.16%
Fold2	5	1	2	70	32	40	43	39	40	11.39%	59.49%	49.37%	45.57%	50.63%	49.37%
Fold3	5	1	2	59	39	36	46	43	41	25.32%	50.63%	54.43%	41.77%	45.57%	48.10%
Fold4	5	1	2	59	37	35	46	43	44	25.32%	53.16%	55.70%	41.77%	45.57%	44.30%
Fold5	5	1	2	41	39	39	35	34	33	48.10%	50.63%	50.63%	55.70%	56.96%	58.23%
Average Error										32.91%	53.67%	51.90%	47.59%	50.63%	50.63%

Table B 26. Feature Selection with Stepwise on PIMA Diabetes Data

FS: Stepwise on PIMA Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	6	1	35	27	35	39	38	38	77.12%	82.35%	77.12%	74.51%	75.16%	75.16%
Fold2	2	6	1	45	34	41	44	39	39	70.59%	77.78%	73.20%	71.24%	74.51%	74.51%
Fold3	2	6	3	41	39	39	43	41	38	73.20%	74.51%	74.51%	71.90%	73.20%	75.16%
Fold4	2	6	1	33	34	36	37	37	36	78.43%	77.78%	76.47%	75.82%	75.82%	76.47%
Fold5	2	6	1	43	39	32	44	40	39	71.90%	74.51%	79.08%	71.24%	73.86%	74.51%
Average Error										74.25%	77.39%	76.08%	72.94%	74.51%	75.16%

Table B 27. Feature Selection with Welch t-Statistics on PIMA Diabetes Data

FS: Welch t-Statistics I on Pima Diabetes Data				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	6	8	33	36	35	31	33	32	78.43%	76.47%	77.12%	79.74%	78.43%	79.08%
Fold2	2	6	8	41	36	34	36	31	33	73.20%	76.47%	77.78%	76.47%	79.74%	78.43%
Fold3	2	6	8	37	36	38	43	41	40	75.82%	76.47%	75.16%	71.90%	73.20%	73.86%
Fold4	2	6	8	38	42	31	38	34	35	75.16%	72.55%	79.74%	75.16%	77.78%	77.12%
Fold5	2	6	8	38	38	33	42	39	39	75.16%	75.16%	78.43%	72.55%	74.51%	74.51%
Average Error										75.56%	75.42%	77.65%	75.16%	76.73%	76.60%

Table B 28. Feature Selection with Fishers Correlation Score on PIMA Diabetes Data

FS: Fisher Correlation Score I on PIMA Diabetes				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	6	8	33	36	35	31	33	32	78.43%	76.47%	77.12%	79.74%	78.43%	79.08%
Fold2	2	6	8	41	36	34	36	31	33	73.20%	76.47%	77.78%	76.47%	79.74%	78.43%
Fold3	2	6	8	37	36	38	43	41	40	75.82%	76.47%	75.16%	71.90%	73.20%	73.86%
Fold4	2	6	8	38	42	31	38	34	35	75.16%	72.55%	79.74%	75.16%	77.78%	77.12%
Fold5	2	6	8	38	38	33	42	39	39	75.16%	75.16%	78.43%	72.55%	74.51%	74.51%
Average Error										75.56%	75.42%	77.65%	75.16%	76.73%	76.60%

Table B 29. Feature Selection with Independently Consistent Expression on PIMA Diabetes Data

FS: Independently consistent Expression on PIMA Diabetes				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	1	5	36	36	40	37	36	37	76.47%	76.47%	73.86%	75.82%	76.47%	75.82%
Fold2	2	1	5	70	32	40	43	39	40	54.25%	79.08%	73.86%	71.90%	74.51%	73.86%
Fold3	2	7	1	54	39	38	43	42	40	64.71%	74.51%	75.16%	71.90%	72.55%	73.86%
Fold4	2	1	5	59	37	35	46	43	44	61.44%	75.82%	77.12%	69.93%	71.90%	71.24%
Fold5	2	5	7	38	45	39	39	37	37	75.16%	70.59%	74.51%	74.51%	75.82%	75.82%
Average Error										66.41%	75.29%	74.90%	72.81%	74.25%	74.12%

Table B 30. Feature Selection with Mean Difference Score on PIMA Diabetes Data

FS: Mean Difference Score on PIMA Diabetes				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	6	8	33	36	35	31	33	32	78.43%	76.47%	77.12%	79.74%	78.43%	79.08%
Fold2	2	6	1	45	34	41	44	39	39	70.59%	77.78%	73.20%	71.24%	74.51%	74.51%
Fold3	2	6	8	37	36	38	43	41	40	75.82%	76.47%	75.16%	71.90%	73.20%	73.86%
Fold4	2	6	8	38	42	31	38	34	35	75.16%	72.55%	79.74%	75.16%	77.78%	77.12%
Fold5	2	6	8	38	38	33	42	39	39	75.16%	75.16%	78.43%	72.55%	74.51%	74.51%
Average Error										75.03%	75.69%	76.73%	74.12%	75.69%	75.82%

Table B 31. Feature Selection with Average Difference Score on PIMA Diabetes Data

FS: Average Difference Score on PIMA Diabetes				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	2	4	7	55	35	39	38	37	36	64.05%	77.12%	74.51%	75.16%	75.82%	76.47%
Fold2	2	8	5	39	35	33	36	38	38	74.51%	77.12%	78.43%	76.47%	75.16%	75.16%
Fold3	2	6	7	39	38	40	38	37	39	74.51%	75.16%	73.86%	75.16%	75.82%	74.51%
Fold4	7	2	8	39	37	35	34	33	32	74.51%	75.82%	77.12%	77.78%	78.43%	79.08%
Fold5	6	8	7	48	43	41	44	46	44	68.63%	71.90%	73.20%	71.24%	69.93%	71.24%
Average Error										71.24%	75.42%	75.42%	75.16%	75.03%	75.29%

Table B 32. Feature Selection with Relief Algorithm on PIMA Diabetes Data

FS: Relief Algorithm on PIMA Diabetes				Error using Equation (4.9)	Error using Equation (4.10)	Error using ANFIS	Error Using Fuzzy KNN (N=5)	Error Using Fuzzy KNN (N=10)	Error Using Fuzzy KNN (N=15)	Accuracy using Eq. (4.9)	Accuracy using Eq. (4.10)	Accuracy using ANFIS	Accuracy using F-KNN (N=5)	Accuracy using F-KNN (N=10)	Accuracy using F-KNN (N=15)
Fold1	8	1	5	71	48	43	52	51	50	53.59%	68.63%	71.90%	66.01%	66.67%	67.32%
Fold2	8	5	1	51	50	48	48	47	47	66.67%	67.32%	68.63%	68.63%	69.28%	69.28%
Fold3	1	8	5	73	47	43	42	41	41	52.29%	69.28%	71.90%	72.55%	73.20%	73.20%
Fold4	8	1	5	69	44	42	44	42	43	54.90%	71.24%	72.55%	71.24%	72.55%	71.90%
Fold5	8	5	1	55	47	48	48	47	48	64.05%	69.28%	68.63%	68.63%	69.28%	68.63%
Average Error										58.30%	69.15%	70.72%	69.41%	70.20%	70.07%

Vita

Sean N. Ghazavi was born in Lafayette, Louisiana on December 25, 1979. He received his Bachelor of Science degree in Industrial and Systems Engineering with minors in Mathematics and Economics from Virginia Polytechnic and State University on December 2004. Upon graduation he worked part time in APV Inc. in Washington DC. Interested in describing system behaviors, making predictions, and optimizing processes he chose Industrial Engineering as his major and enrolled in Louisiana State University on August 2006 to pursue his Masters degree working on data mining and artificial intelligence. The degree of Master of Science will be conferred at the December 2007 commencement.