Chapter 5

A Hybrid Approach of Adaptive Neuro Fuzzy Inference System and Novel Relief Algorithm

---

The objective of this chapter is to create a computer-aided diagnostic model that help in the early detection of breast cancer and hence decrease death rate. The study introduces a hybrid strategy of effectively diagnose the breast cancer by using a novel relief algorithm for feature selection with an Adaptive Neuro-Fuzzy Inference System (ANFIS).

The chapter is organized as follows: Section 5.1 represents the general introduction of ANFIS. Section 5.2 briefly introduces methodology of Adaptive Neuro-Fuzzy Inference System (ANFIS), novel Relief algorithm. Section 5.3 represents experiments, comparative analysis and explanation of the results obtained for WBC data set. The Conclusion is included in Section 5.4.

5.1 Introduction

Systems that are poorly defined or unclear are difficult to model using traditional mathematical techniques (such as differential equations) [55]. Alternatively, the qualitative components of human knowledge and reasoning processes can be modelled by a fuzzy inference system using fuzzy if-then rules without resorting to precise quantitative analysis. First systematically investigated by Takagi and Sugeno [113], fuzzy modelling or fuzzy identification has since found widespread use in control [93] [114], prediction, and inference [56]. However, there are some basic aspects of this approach which are in need of better understanding. More specifically: 1) There are no established protocols for assimilating human expertise into the fuzzy inference system's rule base and data repository and 2) Efficient strategies are required for optimizing the membership functions (MFs) to either reduce the output error measure or increase the performance index.

Medical diagnostic decision assistance systems have grown in prominence as a vital part of modern medicine. The fundamental principle of modern medical technology is an inductive engine that learns the characteristics of human decision-making disorders and can be used to diagnose individuals in the future with varying degrees of illness certainty. Multilayer perceptron neural networks (MLPNNs), Convolutional Neural Networks (CNNs), probabilistic neural networks (PNNs), recurrent neural networks (RNNs), and support vector machines (SVMs) are utilised in medical diagnostic assistance systems to aid human decision-makers in disease diagnosis. The fields of information technology, production technique, decision making, pattern recognition, diagnostics, data analysis, etc. are increasingly interested in and dependent on fuzzy sets. Fuzzy systems that employ ANNs theory to learn about their attributes (fuzzy sets and fuzzy rules) through processing data samples are called neuro-fuzzy systems.

According to the article by L. A. Zadeh , "Knowledge-based systems are based on fuzzy logic which has been applied in many fields like home appliances, automobiles, control, medicine" [131]. Fuzzy Logic employs its lexicon, such as Fuzzification, Defuzzification, Membership Function, Linguistic variables, Domain, Rules, and so on [131]. On the core of human intelligence, knowledge, and understanding, strategies are developed by the Fuzzy Logic Control System (FLC) to handle forbidding operations. Nevertheless, expert systems cannot manage complex processes, whereas FLC systems are utilized for ambiguous processes [16]. ANN has learning capabilities, but it cannot interpret results. That is, it acts as a black box. Whereas Fuzzy

Inference System (FIS) can interpret the results using a rule base, it cannot learn. To overcome the disadvantages of both the techniques, joint use of NN and FIS can be employed to get better results. Feature selection also plays a vital role in machine learning. Machine learning starts by adding features as many features as possible to achieve more reliable outcomes. However, the model's performance degrades as the number of features increases, also known as "The Curse of Dimensionality." Due to this system do not behave appropriately and accurately, which is a significant challenge. To deal with such a problem, we propose a novel Relief algorithm for feature selection. Advantages of both artificial neural networks and fuzzy inference systems have been merged into the ANFIS. The ANFIS's benefits include its nonlinear process capture, flexibility in adaptation, and speed of learning. The diagnosis of breast cancer carries with it significant financial and societal repercussions. As a direct consequence of this, several scholars are working in the field of computational intelligence.

P.R. Innocent et al. conducted a study of fuzzy methods for medical diagnosis in nursing assessment using Type-II fuzzy sets [51]. Faran Blag et al. designed a control system using fuzzy logic for the normality of human function in the human brain and also made a medical diagnosis of brain tumor and hemorrhage [16]. Manish Rana et al. proposed an expert system using fuzzy logic to diagnose hemorrhage, brain tumor, cardiac disease, and thyroid [98]. J. B. Awotunde et al. proposed a medical diagnosis system using fuzzy logic for malaria disease [15]. M. A. Madkour et al. developed a model using Fuzzy logic for the diagnosis of Flu [73]. They had also implemented this model for common Measles, German measles, Mumps, Chickenpox, Whooping cough, Common cold and Meningitis. Elif Derya Ubeyli proposed an integrated view of ANFIS to detect Breast Cancer and tested the model on the WBCD data set [122]. Seyedesh S. N. et al. designed a hierarchical fuzzy neural system with Extended Kalman Filter (EKF) [86]. This model is applied to the WBCD data set. The model is compared with Hierarchical Fuzzy Neural System (HFFN)+ EKF and FNN and obtained the results. Gerald S. et al. proposed a hybrid cost-sensitive fuzzy classification for Breast Cancer diagnosis [108]. They investigated with Michigan, and Diffs burgh style approaches with hybrid GA-ANFIS classifier. Somayesh N. et al. [9] proposed a classification model for Breast Cancer based on advanced multi-dimensional fuzzy NN [86]. They applied Hierarchical FNN + Fuzzy Gaussian Potential NN on WBCD data set with new training algorithm HFNN which use lesser rules and parameters to model the nonlinear system. In the antecedent, they used Gaussian Potential Function as membership function (MF). Shweta saxena et.al. surveyed different machine learning algorithms like AN-

FIS, SVM, fuzzy ANFIS and obtained comparative results with other author work [107]. Manisha Arora et al. experimented Neuro-fuzzy expert system for the WBCD dataset with different MFs and Sugeno-Mamdani fuzzy interface system in MATLAB [12]. Fatima et al. experimented neuro-fuzzy model on the WBCD dataset for recognition of breast cancer [37]. Payam et al. proposed an optimized ANFIS model with a feature selection method for detecting breast cancer at an early stage [132]. In the study, they used Association Rules (AR) method to select features from the dataset. In the proposed model Cuckoo optimization algorithm is used to learn ANFIS and has obtained high classification accuracy.

Ejiofor et al. used the ANFIS model for the diagnosis of Breast Cancer using MATLAB [35]. A Sakthviel et al. compared SVM and ANFIS for breast Cancer detection [13]. In the study, they used the Rough and Genetic algorithm as a feature selection technique. The Classification accuracy of both techniques is compared. Indu Bala et al. implemented fuzzy classification with comprehensive learning Gravitational search algorithm in Brain Tumor detection. Also, they used 10-fold cross-validation to split the train-test data set [17]. Ahmed Rizal et al. developed ANFIS with subtractive clustering in breast cancer detection and reduced error of classification [100]. Amany M. L. et al. proposed two approaches in the diagnosis of Breast Cancer [82]. In the first approach, they used the evolution Genetic Algorithm with ANFIS and obtained results. In the second approach, ANFIS is implemented with Principal Component Analysis as a feature selection technique and obtained comparative results. Both proposed models are compared for accuracy. Indira Muhic introduced a new approach to Breast Cancer diagnosis using a Fuzzy c-means algorithm and pattern recognition method [83]. This method is applied to Breast Cancer climate samples and obtained high accuracy. M. Ashraf et al. introduced an information gain technique with ANFIS for Breast Cancer diagnosis [14]. Hazlina et al. [19] presented the ANFIS model for Breast Cancer survival [43]. In these studies, they used a partial logistic ANN model to predict the hazard curve and survival wave of Breast Cancer patients . Wahyuni Eka Sari et al. presented a comparative study on fuzzy Madami-Sugno Tsukamoto for the children's tuberculosis diagnosis [106].

In this chapter, we modify the existing Relief algorithm for feature selection by eliminating outliers, imputing missing values and using the Mahalanobis distance technique. We comprehend this modified relief algorithm as novel relief algorithm. We propose two ANFIS models with two approaches. One approach uses a novel Relief algorithm and the other approach do not incorporate any feature selection technique. Both the ANFIS models have been tested on Wisconsin Breast Cancer

Data set (WBCD) data set and classification accuracy is obtained. The results shows that the hybrid simulation of the Fuzzy Inference System with ANN and novel Relief algorithm gives the highest 99.30% classification accuracy.

5.2 Methodology

5.2.1 Adaptive Neuro-Fuzzy Inference System

The ANFIS is developed by Jyh-Shing Roger Jang in 1993 [55]. Adaptive Neuro-Fuzzy Interface System is a hybrid of Artificial Neural Network and Fuzzy Interface system, also known as adaptive network-based fuzzy interface system. ANFIS is a set of algorithms of ANN that is modeled on the Takagi-Sugeno Fuzzy Interface system. An adaptive network is a multilayer feed-forward network made up of nodes connected by directed interconnections. Each node executes the specific function on its receiving signals to produce a signal node output [55] [126]. In the adaptive network, each interconnection describes the direction of signal flow from one node to another, where no weight is assigned to the network. The adaptive network runs for two types of nodes, i.e., static node and adaptive node. Figure 5.1 represents the Architecture of ANFIS model with two inputs, one output, and two rules. The overall system design is made up of five layers, namely i) Fuzzification layer, ii) Rule layer, iii) Normalized layer, iv) Defuzzification layer, v) Overall output (or summation neuron).

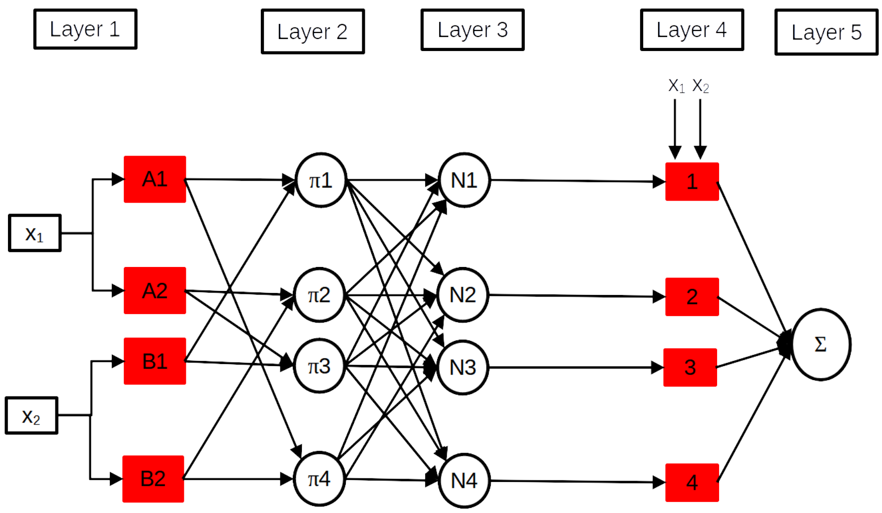

Figure 5.1: Architecture of ANFIS model with two inputs, one output, and two rules

A basic Takagi-Sugeno fuzzy rule is describe by:

$$\text{IF}(x_1 is A_1)\text{AND}(x_2 is A_2 ...\text{AND}(X_m is A_m)\text{THEN} y = f(x_1, x_2, ..., x_m)$$

Where, $x_1, x_2, ..., x_m$ are input variables; $A_1, A_2, ..., A_m$ are fuzzy sets and y is either a constant or linear function of input variables [55] [113].

Consider the first order, Takagi-Sugeno FIS, for two rules [126] [87].

$$\text{Rule 1 : IF}(x_1 is A_1)\text{AND}(x_2 is B_1)\text{THEN} y_2 = f_1 = k_{10} + k_{11}x_1 + k_{12}x_2$$
$$\text{Rule 2 : IF}(x_1 is A_2)\text{AND}(x_2 is B_2)\text{THEN} y_2 = f_2 = k_{20} + k_{21}x_1 + k_{22}x_2$$

Where, x_1 and x_2 are input variables. A_i and B_i are fuzzy set which represents the membership function of ANFIS antecedent. $k_{i0}, k_{i1} and k_{i2}$ are linear consequent parameters which are specified for each rule i. As shown in fig. 5.1, a circle indicates the static or fixed node and a square indicates the adaptive node, i.e., during adaption or training, the parameter can be novel. A brief overview of the ANFIS algorithm is explained as follows [126] [87]:

1. **Fuzzification layer:**
   In this layer, each crisp input is fuzzified using the membership function. All the nodes in this layer are adaptive nodes.

$$O_i^1 = \mu_{A_i}(X); i = 1, 2 \tag{5.1}$$

$$O_i^1 = \mu_{B_j}(X); j = 1, 2 \tag{5.2}$$

   Where, $\mu_{A_i}(X)$ and $\mu_{B_j}(X)$ can be determined using any membership function. Let Gaussian membership function is employed to fuzzify the input variables. The Gaussian membership function is given by eq. 5.3,

$$\mu_{A_i}(x_1) = \exp\left\{-\left(\frac{x_1 - a_i}{b_i}\right)^2\right\}; i = 1, 2 \tag{5.3}$$

   Where a_i and b_i are parameters of membership functions.

2. **Rule Layer:**
   The node of this layer are fixed nodes. They are labeled with π. Each neuron of this layer is associated with a particular Sugeno-FIS and calculates the

firing strength or weight of the rules, respectively. The output of this layer is obtained as follows in eq. 5.4:

$$O_i^2 = w_i = \mu_{A_i}(x_1)\mu_{B_j}(x_2); i = 1,2 \tag{5.4}$$

where, w_i represents the firing strength of the i^th rule.

3. **Normalized layer:**
   Nodes of this layer are also fixed nodes. In this layer, each node represents the normalized firing strength which is obtained in the previous layer and it is indicated by N, as shown in fig. 5.1. The normalization of the firing strength is found using the following eq. 5.5.

$$O_i^3 = \bar{w_l} = \frac{w_i}{w_1 + w_2}; i = 1,2 \tag{5.5}$$

4. **Defuzzification layer:**
   The nodes of this layer are adaptive nodes. The output of this layer is obtained using weighted average defuzzification. The nodes receive initial inputs x_1 and x_2. The resulting output of this layer is computed using the following eq. 5.6

$$O_i^4 = \bar{w_l} f_i = \bar{w_l}(k_{i0} + k_{i1}x_1 + k_{i2}x_2); i = 1,2 \tag{5.6}$$

5. **Overall output:**
   This layer is referred to as the output layer since it aggregates all the nodes of the previous layer 4 and converts fuzzy classification results into crisp values. This layer has a single fixed node labeled as $\sum$. Overall output can be calculated using eq. 5.7

$$O_i^5 = \sum_{i=1}^{2} \bar{w_l} f_i = \frac{\sum_{i=1}^{2} w_i f_i}{w_1 + w_2} \tag{5.7}$$

The first and the fourth layer of ANFIS are adaptive layers. Parameters a_i and b_i of membership functions at layer one are adaptable, known as premise (or antecedent) parameters. Parameters k_{i0}, k_{i1}, k_{i2} of layer four are adaptive and known as consequent parameters.

The learning algorithm of ANFIS

Premise parameter namely $\{a_i, b_i\}$ and consequent parameter $\{k_{i0}, k_{i1}, k_{i2}\}$ are to be learn to calculate the output of ANFIS for given training data. The output of the ANFIS can be written as in eq. 5.8:

$$\begin{aligned} f &= \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2 \\ &= \bar{w_1}(k_{10} + k_{11}x_1 + k_{12}x_2) + \bar{w_2}(k_{20} + k_{21}x_1 + k_{22}x_2) \\ &= (\bar{w_1})k_{10} + (\bar{w_1}x_1)k_{11} + (\bar{w_1}x_2)k_{12} + (\bar{w_0})k_{20} + (\bar{w_2}x_1)k_{21} + (\bar{w_2}x_2)k_{22} \end{aligned} \tag{5.8}$$

Equation 5.8 represents the linear combination of adaptable consequent parameters k_{i0}, k_{i1} and k_{i2}. These consequent parameters are obtained using the Least Square method [87].

ANFIS employs a hybrid learning technique that combines the least-squares method with the gradient descent approach. ANFIS is composed of the forward pass and the backward pass as shown in table 5.1. Once optimal consequent parameters are obtained, premise parameters are adjusted using the Gradient Descent method.

Table 5.1: Two passes to learn parameter for ANFIS

	Forward Pass	Backward Pass
Premise Parameter	Fixed	Gradient Descent method
Consequent Parameter	Least-Square method	Fixed

The fundamental goal of the ANFIS is to use a learning algorithm to obtain the optimal values of the parameters.

5.2.2 Novel Relief algorithm

The Relief algorithm is a feature selection technique. The Relief algorithm was developed by Kira & Rendall in 1992 which was influenced by instance-based learning [62]. This algorithm is capable of dealing with both nominal and numerical variables. It cannot, however, cope with missing data and outliers and it is constrained to the two-class problem only. To overcome these problems, we propose new distance-based learning Relief algorithm and comprehend as novel relief algorithm. The proposed novel Relief algorithm deals with missing data and outliers.

Features are selected using the Mahalanobis distance technique. Missing values are obtained using Multivariate Imputation by Chained Equation (MICE) and Outliers are detected using Euclidean Distance. Outliers increase data uncertainty and reduce predictive capacity. As a consequence, eliminating outliers will increase the significance of data.

Multivariate Imputation by Chained Equation (MICE)

1. Input all missing values with the mean of their respective columns as a starting point. Call it a 'zeroth' dataset. Input columns from left to right.

2. Remove the imputed values (from the 1st column from left), say x.

3. The remaining feature and rows become the feature matrix and $'x_1'$ becomes the target variable. Apply Multivariate Linear Regression (MLR) model on the rest of the columns $'x_2, x_3, ..., x_n'$ to estimate the missing value in x_1 (as test data) and replace the missing value with the newly computed value. Repeat this procedure for those columns that have missing values.

4. Then subtract 'zeroth data set' with the latest "transformed" data set from which this new data set was formed. Repeat this step 2 and 3 with this unique data set, until getting the stable model, i.e., until the difference between the two latest imputed data sets becomes very small.

5. Stop iterations when a pre-defined threshold is breached or do it until a pre-defined maximum number of iterations get completed.

Algorithm

Consider the training samples of size n which is represented by $X = (x_1, x_2, ..., x_n)$. Each of these samples consists of a features which are given by $A = \{A_1, A_2, ..., A_a\}$. Each sample has a target value O_j. τ is the threshold parameter. The algorithm is briefly explained here.

- Input: n number of training samples having a vector of feature values and class values respectively.

- Output: The vector of W of feature scores which estimates the relevance of features.

1. Find missing values using MICE.
   Find outliers of data using Euclidean distance.
   Select m number of random training samples from n number of training samples ($n > m$).

2. Initialize all feature weights. say $W[A] = 0.0$

3. For $i = 1$ to m do begin
   Choose random 'target' instance R_i.
   Find nearest hit $'H'$ (of the same class value, say +ve) and nearest miss $'M'$ (of the opposite class value of the sample, say -ve using Mahalonbis Distance.
   $d(x, y) = \sqrt{(\overrightarrow{x} - \overrightarrow{\mu})^2 s^{-1} (\overrightarrow{x} - \overrightarrow{\mu})}$
   where, $\mu \rightarrow$ mean and $s \rightarrow$ covariance matrix.

4. For each feature, update the weight:
   for $A = 1$ to a do
   $W[A] = W[A] - diff(A, R_i, H_i) + diff(A, R_i, M_i)$
   Where, $diff(A, R_i, H_i) = \frac{\sqrt{(R_i - H_i)^2 s^{-1} (R_i - H_i)}}{maxR_i - minR_i}$ and
   $diff(A, R_i, M_i) = \frac{\sqrt{(R_i - M_i)^2 s^{-1} (R_i - M_i)}}{maxR_i - minR_i}$

5. Find relevance of feature: $R_{fi} = \frac{W}{m}$

6. For $i = 1$ to a , if $(R_{fi} \geq \tau)$ then The feature f_i is relevant else The feature f_i is irrelevant.

7. Print the vector W having feature performance with relevance.

In the second phase of this algorithm, we find the vector W having feature performance using Mahalanobis distance. The novel Relief algorithm iteratively loops through m random training samples ($T_(R_i)$) out of n samples without repetitions. The weight of the feature can range from -1 to +1, where -1 indicates the worst score and +1 indicates the best score. Finally, a novel Relief algorithm selects those features whose average weights are greater than the defined threshold.

5.3 Experiments

The performance of the novel Relief-based ANFIS classifier is evaluated on the 'Wisconsin Breast Cancer Data set (WBCD)' for breast cancer classification. Four features, namely Clump Thickness, Uniformity of cell size, Marginal Adhesion and

Mitoses, are found significant using the novel Relief algorithm. Each feature consists of 3 membership functions and hence produces $3^4 = 81$ fuzzy rules.

The proposed ANFIS model with the novel Relief algorithm is trained and tested as depicted in fig. 5.2. In the first step, we input data to the ANFIS model. Then in the second step, data is normalized. Then a novel Relief algorithm for feature selection is applied to the data set to reduce training time and obtaining better precision with reduced overfitting. As a part of the pre-processing, outliers of the data are found using Euclidean Distance and are neglected and also the missing values are computed using the MICE algorithm.

Figure 5.3 shows the outliers of the WBCD data set, which interpret those areas that are far from the rest of the findings. As a consequence, the analysis performed on the WBCD data set may be biased. Hence, outliers detection is essential and is the first stage of the novel Relief algorithm.

Figure 5.4 depicts the features that are selected if their relevance level is greater than or equal to τ, while the rest are discarded. The results of the relevant feature using the novel relief algorithm are computed for $m = 475$ and $\tau = 0.5$. We observe that from all the features, only 4 features have relevance scores greater than τ. The averages of the 475 runs were used to produce the relevance score. The advantage of using Mahalanobis distance is that it uses group means and variance for each feature and hence eliminating the problems of scale and correlation that are inherent in the Euclidean distance. Using this novel algorithm we find covariance among the features. Redundant information from the highly correlated features are also removed using this algorithm. Hence, we claim that the novel Relief algorithm can deal with missing values, outliers and redundant features.

This algorithm also detects missing values from the data set during the feature selection process while finding the nearest hit and nearest miss using the MICE technique. Out of 699 records of the WBCD data set, 16 records contain missing feature values. For all the records, first the novel relief algorithm finds missing values using the MICE technique, which would be helpful during feature selection.

In the next step of the proposed model, the input-output variables are determined from the WBCD data set based on a novel Relief algorithm. The divergence of the data from their nominal values shows the symptoms of the disease. Fuzzy sets, fuzzy domain, fuzzy rules and fuzzy membership functions are defined using input variables.

In this proposed model, all crisp input values are converted into fuzzy input values.

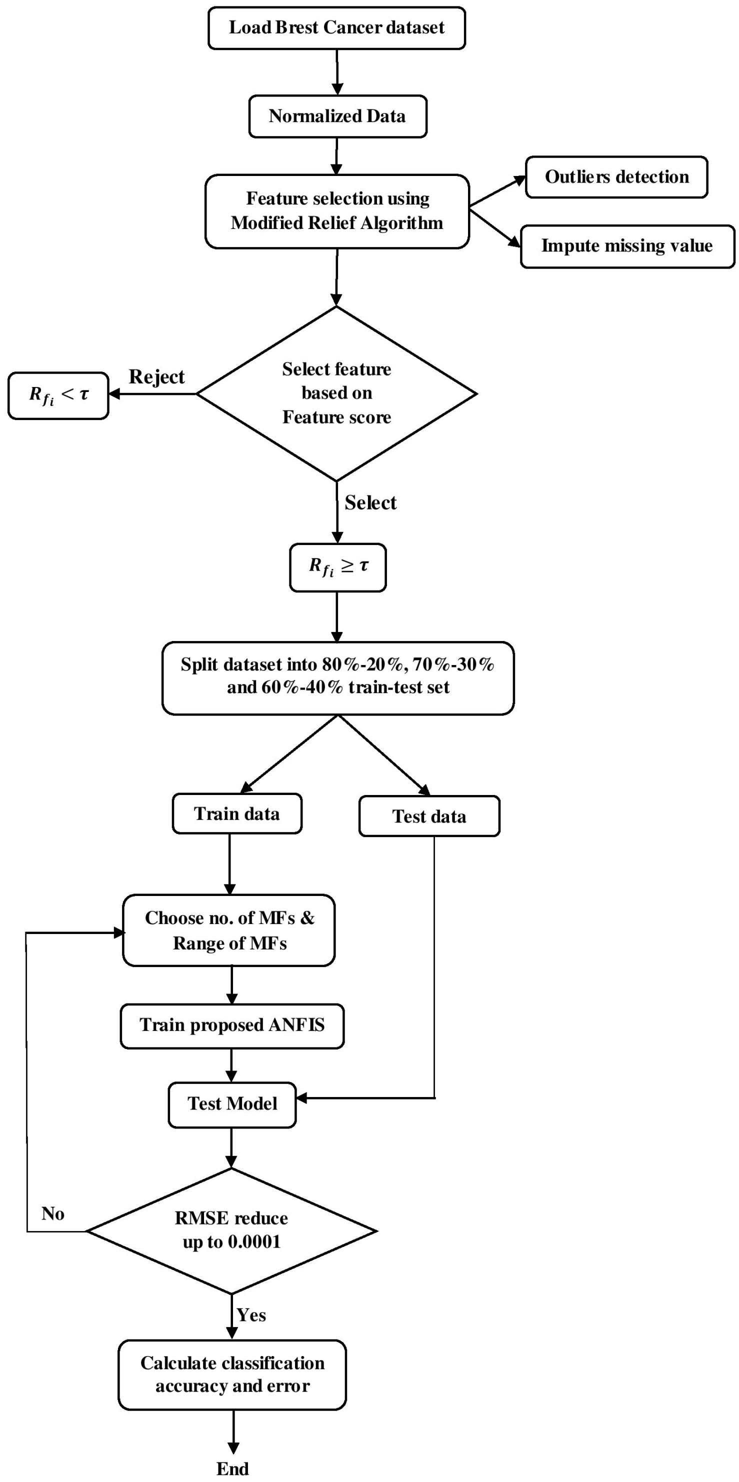

Figure 5.2: Flowchart of proposed ANFIS model

We used the Gaussian membership function to convert crisp values into a fuzzy value which is the first step of fuzzifying the proposed model.

Table 5.2 shows the range of the Gaussian membership function of the four input

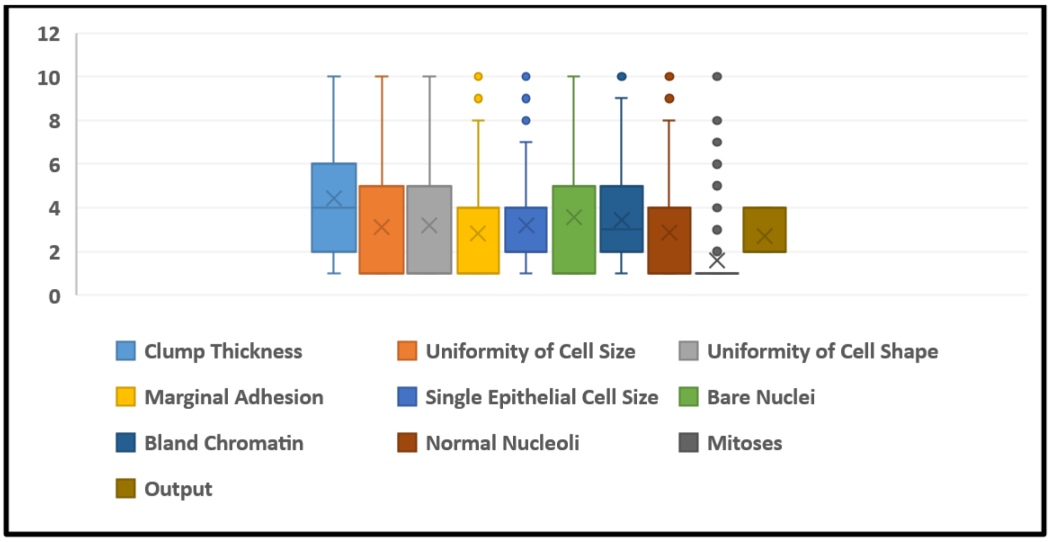

Figure 5.3: Outliers of WBCD data set

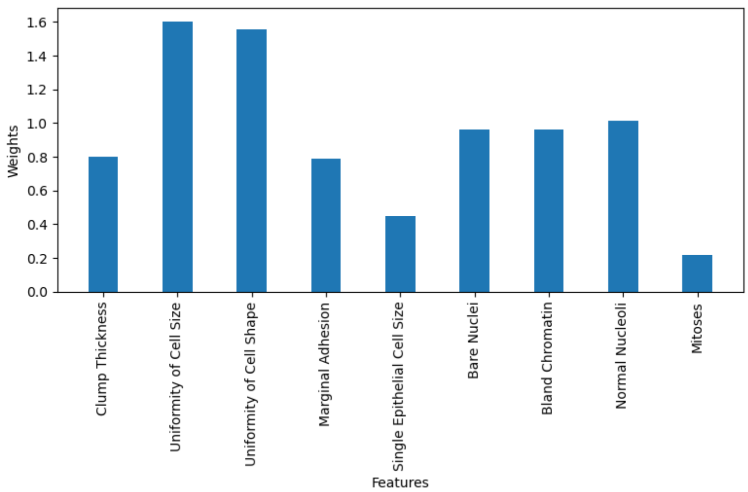

Figure 5.4: Feature Selection using Modified Relief algorithm

features.

Table 5.2: Range of Gaussian membership function for each selected feature

Range of Gaussian membership function	Feature 1	Feature 2	Feature 3	Feature 4
	[-1.214 1.983]	[-0.6995 2.5]	[-0.3437 4.904]	[-0.6332 2.521]

During this fuzzification process, linguistic variables are assessed using the Gaussian membership function and they are represented by an appropriate degree of membership range from 0 to 1 using equation 3 and it is expressed in table 5.3.

Figure 5.5 is the architecture of the Sugeno Fuzzy Inference System for the WBCD

Table 5.3: Gaussian membership function with a range of Linguistic variables for each selected feature

Membership function	Feature 1	Feature 2	Feature 3	Feature 4
Low (L)	[0.6787 -1.214]	[0.6262 -0.6995]	[1.114 -0.3437]	[0.6697 -0.6332]
Medium (M)	[0.6787 0.3844]	[0.6262 0.7752]	[1.114 2.28]	[0.6697 0.9439]
High (H)	[0.6787 1.983]	[0.6262 2.25]	[-0.3437 4.904]	[1.114 4.904]

data set.

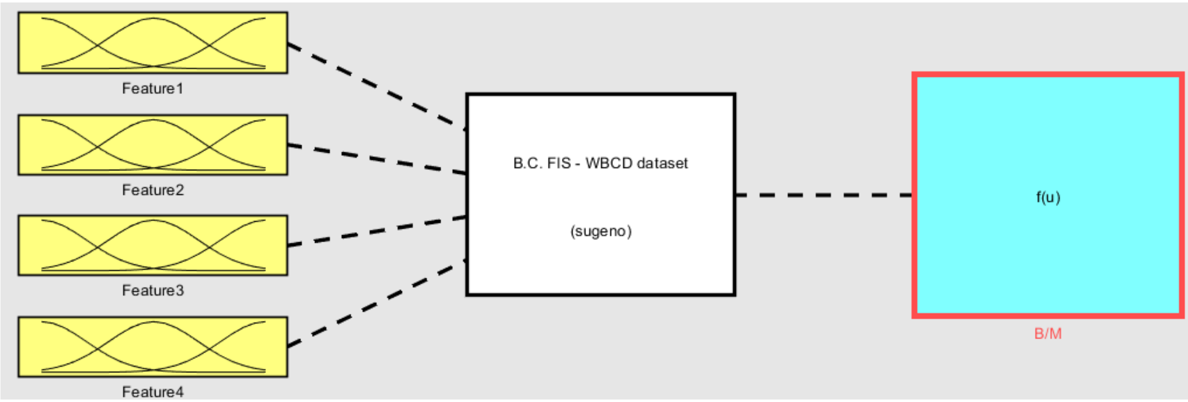

Figure 5.5: Sugeno FIS for WBCD dataset

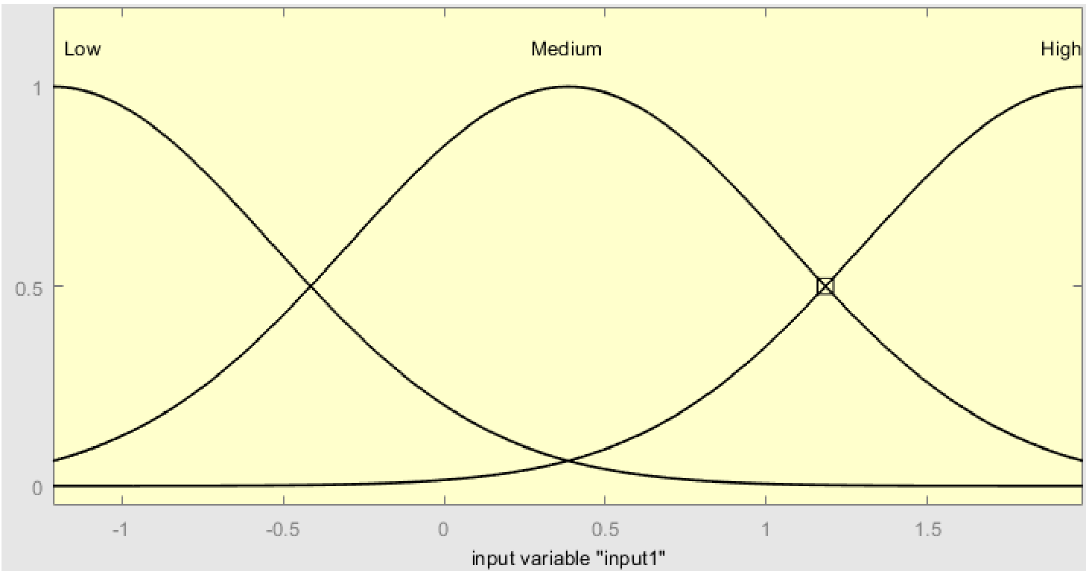

Figure 5.6: Gaussian membership function of Feature 1

Figures 5.6, fig. 5.7, fig. 5.8 and fig. 5.9 depict the Gaussian membership function for each fuzzy feature input, namely features 1, 2, 3, and 4.

The next step is developing fuzzy rules. The proposed model comprises 81 rules out of which only valid rules are claimed to be fired by the domain. The weighted

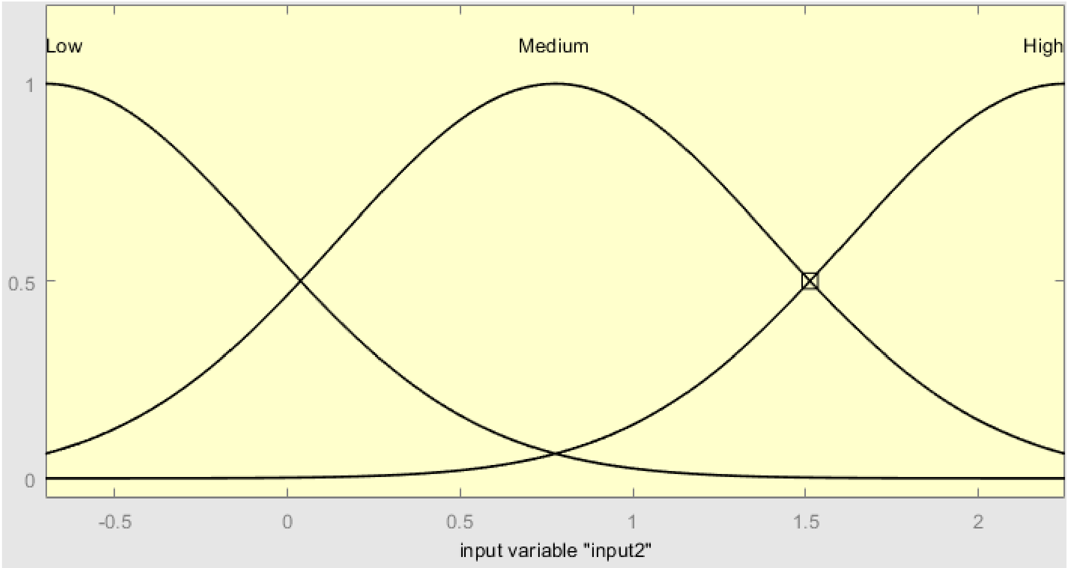

Figure 5.7: Gaussian membership function of Feature 2

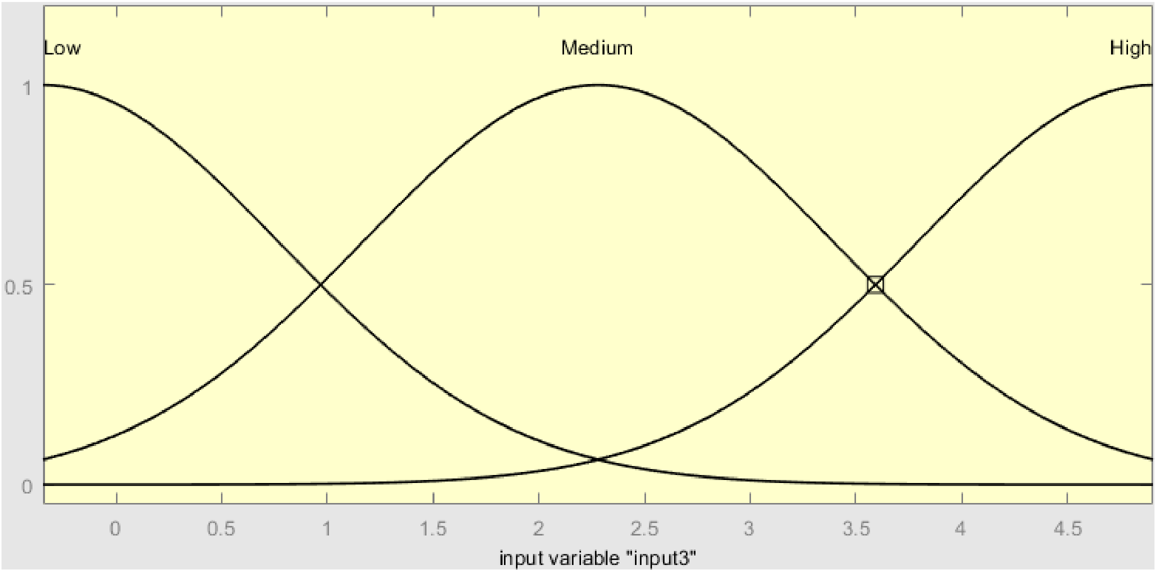

Figure 5.8: Gaussian membership function of Feature 3

average approach is used to defuzzify the fuzzy output into crisp output.

Figure 5.10exhibits the structure of the proposed ANFIS model, which has four feature inputs and one output, which is the ANFIS model's Neural Network representation.

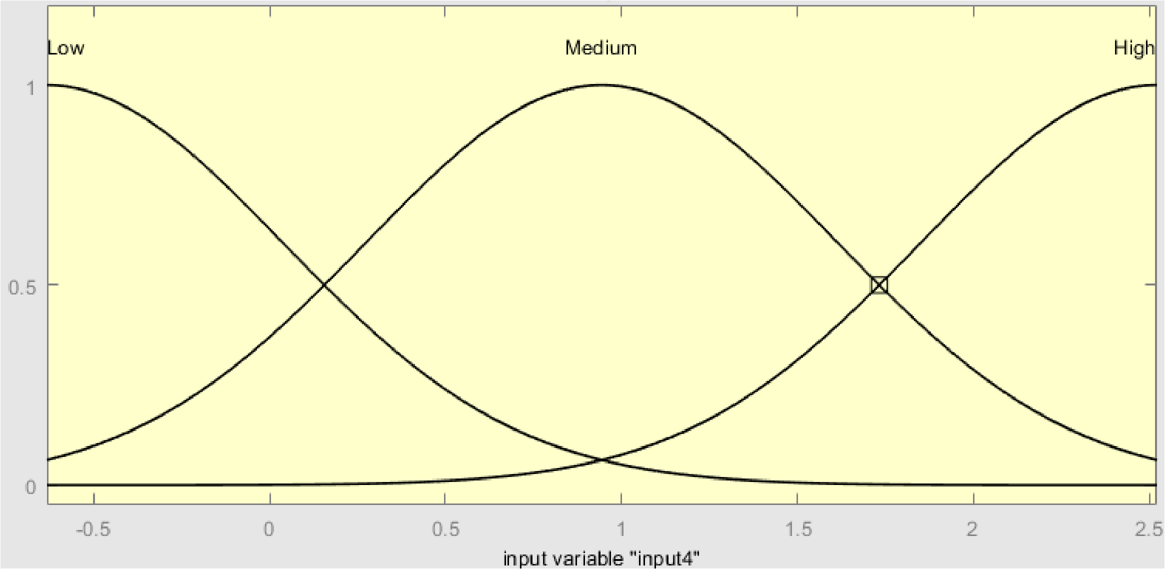

Figure 5.9: Gaussian membership function of Feature 4

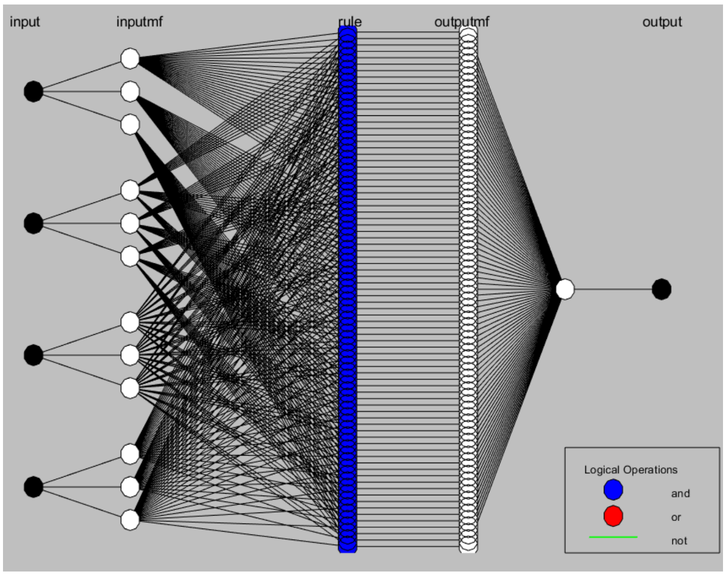

Figure 5.10: Structure of proposed ANFIS model

5.3.1 Results

The model's performance and efficiency are analyzed by accuracy, sensitivity, precision and F-score using the confusion matrix.

The WBCD data set is divided into 80-20%, 70-30%, and 60-40% train-test sets on the basic ANFIS model and the proposed hybrid ANFIS model. Figures 5.11, FIG. 5.12 and FIG. 5.13 demonstrate the effectiveness of a basic ANFIS model with diverse train-test sets. The performance of the Gaussian M.F. model has been further assessed using various learning rates and epochs. We attained the highest classification accuracy of 90.71% with 60-40% train-test set for 0.01 learning rate and 100 epochs with a simple model.

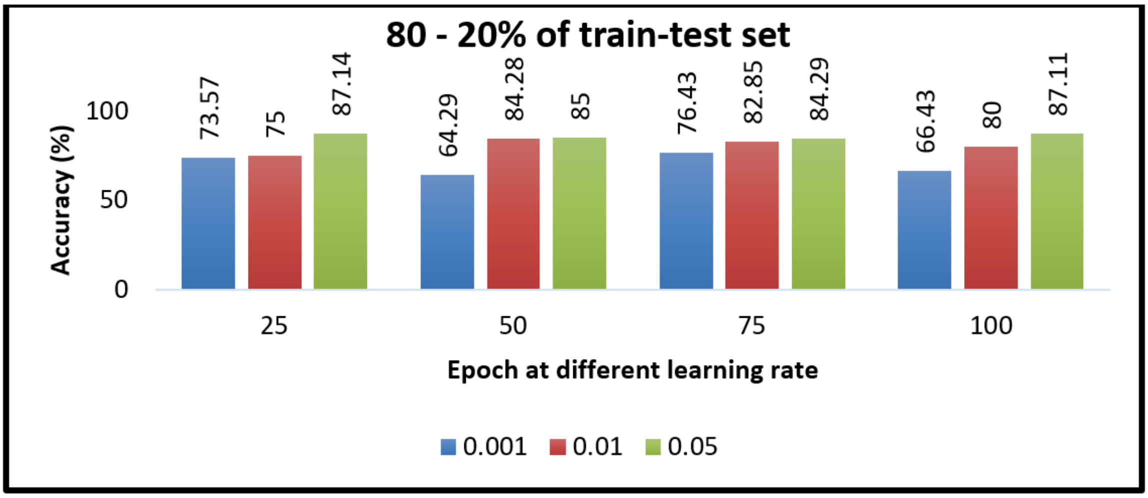

Figure 5.11: Comparison of Accuracy for Simple model: different learning rate with different epoch for 80 - 20% train-test set

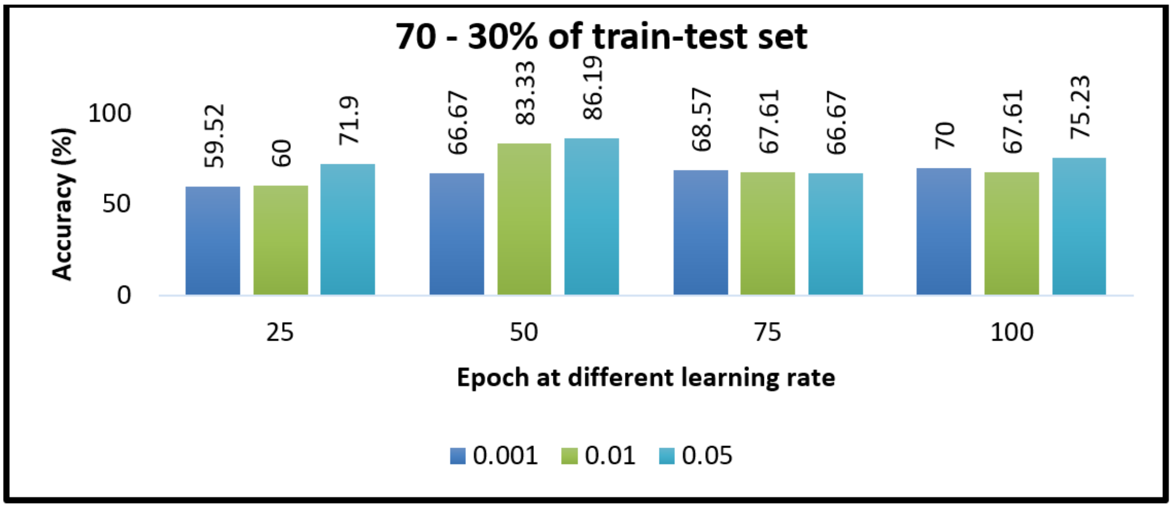

Figure 5.12: Comparison of Accuracy for Simple model: different learning rate with different epoch for 70 - 30% train-test set

Figures 5.14, FIG. 5.15 and FIG. 5.16 illustrate the efficacy of a proposed hybrid ANFIS model using the train-test sets having same ratios as simple ANFIS model. The performance of the Gaussian M.F. model was further evaluated using different learning rates and epochs. The hybid ANFIS model attained the highest classification accuracy of 99.30% using 0.01 learning rate and 50 epochs.

Figure 5.17 depicts the train data set's confusion matrix, which yields a classification rate of 93.20%. It reveals that 1.25% of malignant cases are misdiagnosed and 5.55% of benign cases are misdiagnosed.

Figure 5.18 depicts the confusion matrix for the testing data set, demonstrating that

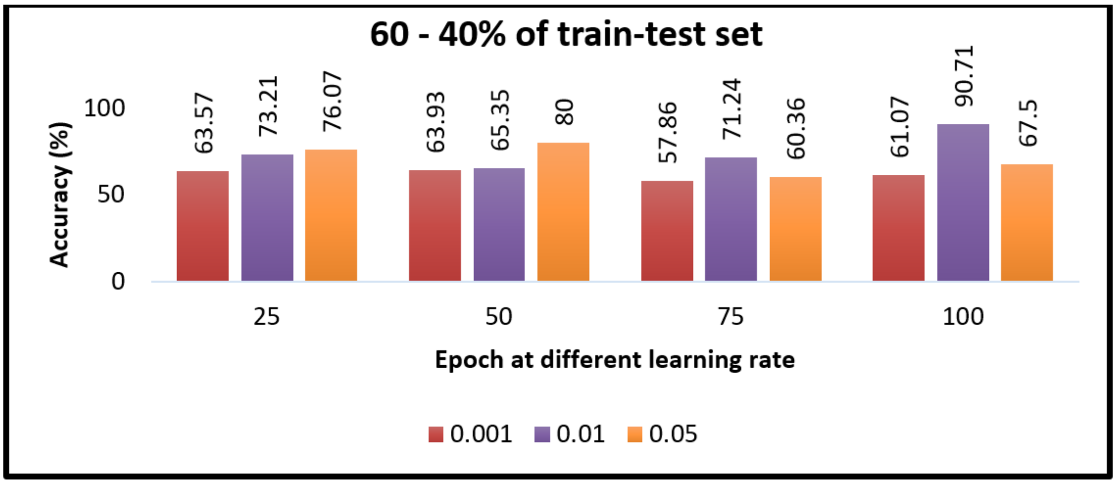

Figure 5.13: Comparison of Accuracy for Simple model: different learning rate with different epoch for 60 - 40% train-test set

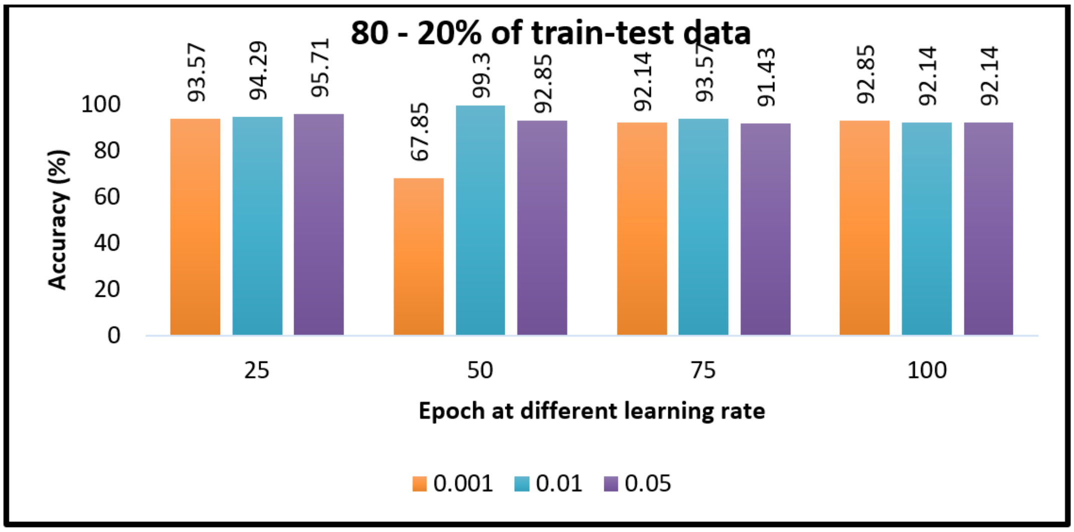

Figure 5.14: Comparison of Accuracy for proposed modified model: different learning rate with different epoch for 80 - 20% train-test set

the proposed model obtained 99.30% classification accuracy with a 0.7% misclassification rate.

The error vs. epoch plot is shown in fig. 5.19, where the error is minimized to 0.001 after a certain epoch for a learning rate of 0.001.

Figure 5.20 depicts the testing performance matched with the actual result. The alignment of each testing data with actual data emphasizes the testing precision.

Table 5.4 depicts the performance analysis and efficiency of the proposed hybrid ANFIS model to the findings of other authors.

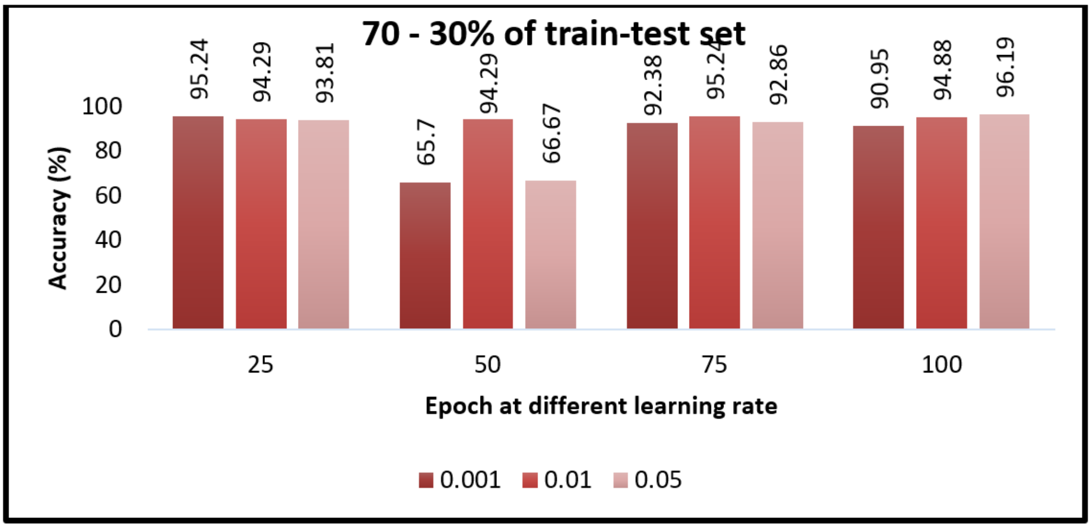

Figure 5.15: Comparison of Accuracy for proposed modified model: different learning rate with different epoch for 70 - 30% train-test set

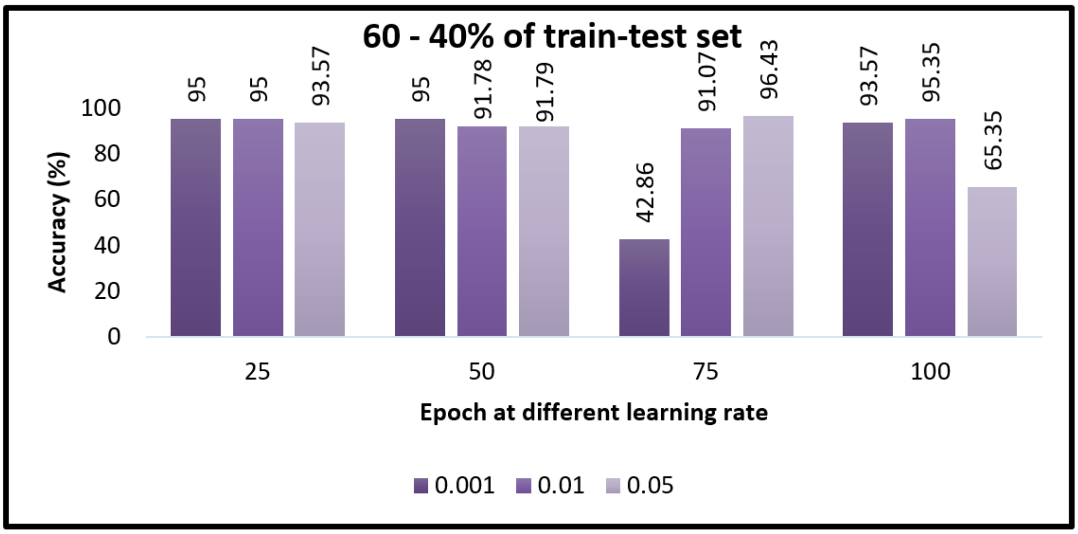

Figure 5.16: Comparison of Accuracy for proposed modified model: different learning rate with different epoch for 60 - 40% train-test set

The detailed information of the both data set are given in the Appendix. Also, above all computation is carried out using Python programming and it is given in the Appendix.

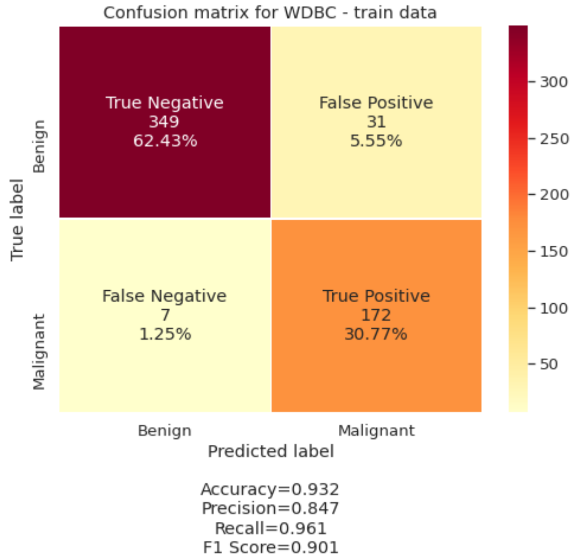

Figure 5.17: Confusion matrix for Train data

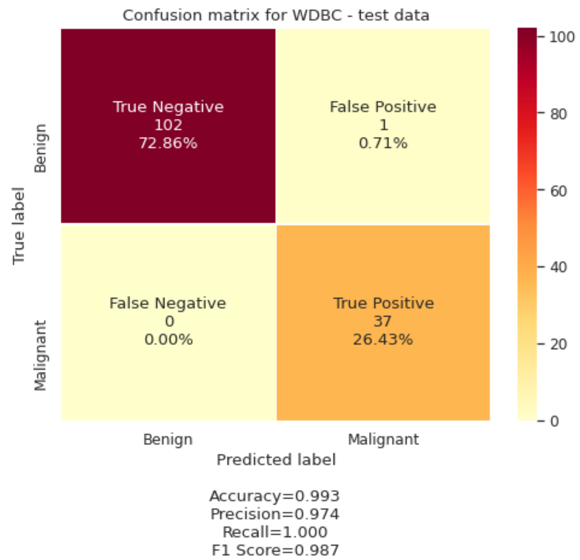

Figure 5.18: Confusion matrix for Train data

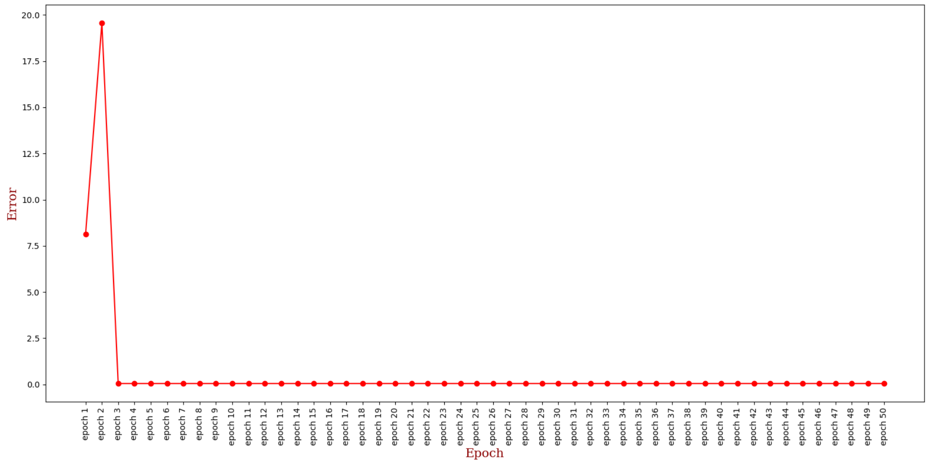

Figure 5.19: Epoch vs. Error

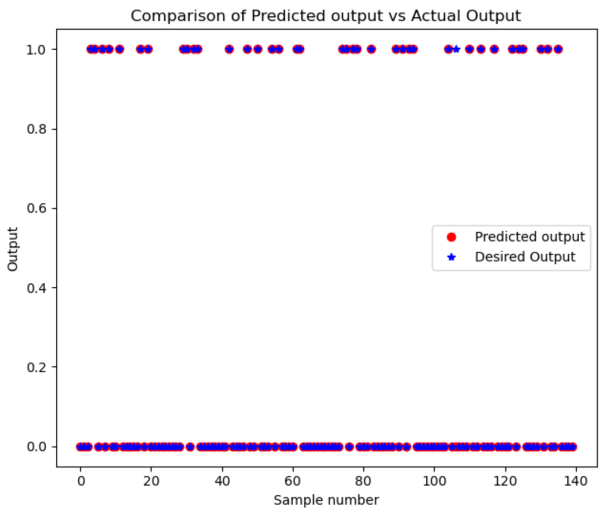

Figure 5.20: Testing performance with actual output

Table 5.4: Performance analysis of most popular Breast Cancer detection methods

Authors	Year	Methodology	Accuracy (%)
Elif Derya Ubeyli	2008	ANFIS	99.10
Gerald Schaefer	2010	Genetic algorithm+ANFIS	97.25
Seyedeh someyeh	2010	Hierarchical ANFIS	99.40
Somayeh Naghibi	2011	ANFIS	98.20
Bekaddour Fatima	2012	ANFIS	98.25
Manisha Arrora	2012	ANFIS	98.58
Payam Zarbaksh	2017	Association+ANFIS	99.26
A. Sakthivel	2018	ANFIS	98.92
		SVM	93.02
Indu Bala	2019	Gravitational Search algorithm+ANFIS	96.12
Ahmad Rizal	2020	ANFIS	98.00
Amany Mostafa	2020	ANFIS	99.10
		GFIS	97.70
Proposed work	2021	Modified Relief algorithm+ANFIS	99.30

5.4 Conclusion

In this study, the existing Relief algorithm for feature selection is modified by eliminating outliers, imputing missing values and using the Mahalanobis distance technique. We also propose an ANFIS model with a novel Relief algorithm for feature selection. This proposed hybrid ANFIS model (with novel Relief algorithm as feature selection technique) and simple ANFIS model (without using feature selection technique) have been validated on the Wisconsin Breast Cancer Data set (WBCD). The proposed hybrid ANFIS model having the Neural Network capabilities and the Fuzzy Interface System as a rule-based system with novel Relief algorithm yields excellent outcomes. The findings indicate that the hybrid approach used in the proposed ANFIS model (with novel Relief algorithm as feature selection technique) gives high accuracy, sensitivity and precision.