

# BREAST CANCER DIAGNOSIS BASED ON GENETIC-FUZZY LOGIC AND ANFIS USING WBCD

Amany Mostafa Lotayef<sup>1</sup>, Khaled H. Ibrahim<sup>2</sup> and Rania Ahmed Abul Seoud<sup>3</sup>

<sup>1</sup>Communication and Electrics, Faculty of Engineering, Fayoum University

<sup>2</sup> Assistant Prof of Electrics Faculty of Engineering, Fayoum University

<sup>3</sup> Prof. Digital Signals Faculty of Engineering, Fayoum University

## ABSTRACT

*Breast cancer is conspicuously one of the most common diseases that cause death for women. Besides, it is increasing with high rates. Consequently, Breast Cancer must be discovered in early stages to avoid death or losing part of the body due to late diagnosis. Thus, there are many researches for computerizing breast cancer diagnosis with different techniques. It reduces human decision rate in order to decrease the mortality rate through the disease. Therefore, we have a major motivation for this highly significant work. The primary focus of the research is to produce a multi-model that can predict the diagnosis whether benign (noncancerous) or malignant (cancerous) nature of a tumor with high accuracy using two methods. The first method is a combining of two major methodologies, namely the fuzzy based systems and the evolutionary genetic algorithms (GFIS). The second method intends to an integrated view of implementing an adaptive neuro-fuzzy inference system (ANFIS) with feature selection using principle component analysis (PCA). Wisconsin breast cancer database (WBCD) is applied because it contained records of patients with known diagnosis. The proposed target of this research compares breast cancer diagnosis based on physical characteristics of the tumour between GFIS and ANFIS. GFIS has achieved a high performance with 97.7% however ANFIF has achieved the highest accuracy with 99.1%.*

## KEYWORDS

*Breast Cancer, Fuzzy Inference System, Genetic Fuzzy Inference System, Adaptive Neural Fuzzy Inference System, WBCD.*

## 1. INTRODUCTION

Breast cancer is conspicuously considered the second leading cause of women's death by cancer over worldwide, so the proposed method focuses the attention to early detection of the disease using automated diagnostic systems to help doctors in making diagnostic decisions. Breast cancer statistics state that there are over 2 million new cases in 2018 [1] [15]. The earlier cancer is diagnosed, the higher of the chance is for successful treatment. The automatic medical diagnostic decision is our proposed aim. The main concept of the medical technology is an inductive engine that learns the decision characteristics of the diseases for diagnosing future patients with uncertain disease cases. It is important to mention in the number of quantitative models that are introduced for breast cancer classification using machine-learning algorithms. These models utilize fuzzy rules inference system for GFIS and ANFIS in the medical datasets.

Fatima S. Ahadi [2] introduces an advanced fuzzy inference system whose coefficients are optimized by statistical analysis methods. Besides, the method is implemented a Mamdani-Type fuzzy inference system to predict whether a case is benign or malignant. The proposed method uses pre-processing dataset for more efficiency. The fuzzy inputs are assigned weights using a binary logistic regression algorithm and t-test respectively to get the most significant feature. Feature selection gives higher performance. But Abir Alharbi [3] implements a fuzzy inference

system optimized with genetic algorithm using the database of Saudi Arabian breast cancer diagnosis to improve classification. This achieves the best performance with three rules up to 97%. Mehrbakhsh Nilashi [4] develops a knowledge-based system for the classification of breast cancer disease using clustering method to cluster the data in similar groups. Then, the proposed method uses Classification and Regression Trees (CART) to generate the fuzzy rules to be used for the classification of breast cancer disease in the knowledge-based system of fuzzy rule-based reasoning method. Besides, they depend mainly on two real-world datasets, WBCD and Mammographic mass. They achieve up to 94.1%. Thanh Nguyen [5] proposes a classification model that combines wavelet transformation and interval type-2 fuzzy logic system. Interval type-2 fuzzy logic system consists of fuzzy c-means clustering based unsupervised learning and genetic algorithm based on parameter tuning. These components of logic system have high computational costs. The wavelet transform functions is used for reducing these computational costs with high performance 97.88%. Hossein Ghayoumi Zadeh [6] uses thermal breast images of 200 patient then the features are extracted. ANFIS is applied on the extracted features. Results of this method achieves a sensitivity of 93%. Payam Zorbakhsh [7] designs ANFIS using feature selection. The proposed method is applied on Wisconsin Breast Cancer Database (WBCD) using ANFIS and AR for feature selection with high performance of (99.2%) which more than ANFIS without feature selection. A. Onan [8] uses a model based on three main phases. The three phases are: instance selection phase, feature selection phase and classification phase. The classification accuracy of their method was 99.7151%. Additionally, there are researches that combine GA and ANFIS as GA-ANFIS for automatic diagnosis as in [19]. Hamza [19] uses GA-ANFIS which GA is used as feature selection to find the best attributes for ANFIS. Moreover GA-ANFIS is used for many application. It is used for Estimation of Regional Rainfall Amount [20]. M. Hanefi CALP [20] uses GA-ANFIS for estimating the amount of annual rainfall and to ensure effective management of the water resources and allow some evaluations and preparations according to possible climate changes with error rate as much as possible.

ANFIS is also used for brain tumour Sudipta Roy et al, [10] develops a classifier system with ANFIS for brain tumor tissue characterization that has been explored. The classifier got 98.25% accuracy on the Harvard benchmark dataset. Firstly, Feature extraction is implemented for the image . Feature extraction is a general term for methods of building combinations of the factors to get around these issues whereas depicting the data with sufficient precision

In the proposed paper, Advanced Genetic-Fuzzy Inference System Tool (GFIS) and Adaptive Neural-Fuzzy Inference System (ANFIS) tool will be implemented using Wisconsin Breast Cancer Diagnosis Data (WBCD) [11] for diagnosing breast cancer through the next sections.

Section 2: Data Description

Section 3: METHODOLOGY

Section 3.1: First method (GFIS).

Section 3.2: Second method (ANFIS).

Section 4: Experimental Results.

Section 5: Comparison and Discussion

Section 6: Conclusion.

## **2. DATA DESCRIPTION**

Breast cancer is the appearance for a new lump or mass which may be a regular or irregular edges and sometimes the mass may be soft, tender, or rounded also. The mass is called tumour that may be more likely to be cancer. WBCD is applied because it is a digitized image of a fine

needle aspirate (FNA). WBCD is obtained from the University of California at Irvine (UCI) and Machine Learning Repository [11]. WBCD is a medical dataset which has nine features with known diagnosis. The details of the attributes are listed in Table 1. It contains 699 instances. Each one represents a follow-up data for one breast cancer case. The samples consist of visually assessed nuclear features of fine needle aspirates (FNAs) taken from patients' breasts. Each feature is in the interval from 1 to 10. Low values represent a normal case (benign) and high values represent a most abnormal case (malignant).

The next step is data pre-processing. The outliers in the dataset which are Non-relevant data, is removed because an outliers are far outside the normal for a variable or population and they are 18 cases

Table 1. WBCD dataset

No	Attribute Description	Value of Attributes	Mean	Standard Deviation
1	Clump Thickness (v1)	1-10	4.44	2.82
2	Uniformity of Cell Size (v2)	1-10	3.15	3.07
3	Uniformity of Cell Shape (v3)	1-10	3.22	2.99
4	Marginal Adhesion (v4)	1-10	2.83	2.86
5	Single Epithelial Cell Size (v5)	1-10	3.23	2.22
6	Bare Nuclei (v6)	1-10	3.54	3.64
7	Bland Chromatin (v7)	1-10	3.45	2.45
8	Normal Nucleoli (v8)	1-10	2.87	3.05
9	Mitoses (v9)	1-10	1.60	1.73

### 3. METHODOLOGY

In present paper, FIS is used for Breast Cancer diagnosis because FIS is implemented to learn a computer "human like -reasoning-style". However it is not effective for large space so two methodology are implemented to get the highest accuracy. The first method is Genetic-fuzzy inference system with Binary Logistic Regression. Second method is Adaptive neuro fuzzy inference system with feature selection. The two proposed Methods will be introduced through next sections.

#### 3.1. Proposed Genetic Fuzzy Inference System (First Method)

Genetic Algorithm is used for large and complex searching spaces and it is able to give optimal solutions for numerous diverse problems, therefore genetic-fuzzy algorithms (GFIS) is considered as a modelling optimization process. GFIS is a fuzzy system classification which the fuzzy parameters are optimized by genetic algorithm (GA) as shown in Figure 1. GA is used to maximize classification performance for fuzzy inference system. In the proposed method GA is used to generate efficient membership and rule-base for efficient fuzzy system. The evolutionary algorithm is used to tune the knowledge contained in the fuzzy system by finding optimal membership values. An initial fuzzy system is defined by an expert. Then, the membership function values are encoded in a genome. GA is used to find the highest performance of proposed system because it often overcomes the local-minima problem seen in other gradient descent-based Optimization method.

In the GFIS applications, there are main approaches which are the Michigan approach, the Pittsburgh approach, and the iterative rule learning approach. In the Michigan approach, GA optimizes the rules or the memberships inside the fuzzy inference system. And they are in constant competition for the best action to be proposed for efficient fuzzy system. In the Pittsburgh approach, the GA maintains competition between a populations of candidate fuzzy systems. The main disadvantage of Pittsburgh is the cost, since a population of a complete fuzzy system has to be evaluated each generation. In the iterative rule learning approach, it is same as Michigan approach, each chromosome in population represents a one rule, and only the best individual is considered the remaining chromosomes are discarded

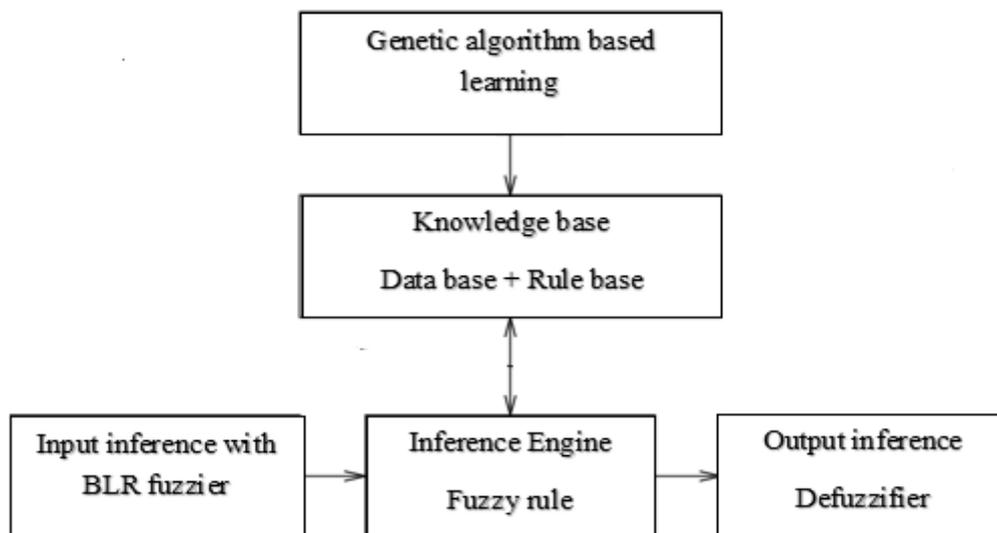


Figure 1. Genetic-Fuzzy Inference System (First Method)

### 3.1.1. Binary Logistic Regression (BLR)

In the proposed method, the data firstly will be pre-processed to improve the accuracy of GFIS. So, BLR will be used to determine the significance of each feature. BLR is performed using IBM SPSS 22 Statistics software because it is used in medical diagnosis. WBCD is normalized then analysed. It includes a dependent variable that can only take on two values. It is a diagnosis and it has two values which are malignant and benign. BLR contains an important information; p-value. P-value depends on (Sig). If the p-value is more than 0.05, then it confirms the null hypothesis and there is no relationship between two phenomena.

### 3.1.2. Proposed Fuzzy Inference System (FIS)

Fuzzy logic is originated by Zadeh [Zadeh65]. The main concept of fuzzy logic examines human thinking and reasoning and mathematically. Fuzzy logic could solve a problems of making a decision. Fuzzy logic is a form of many-valued sets or fuzzy sets i.e. In the fact, FIS classification has an advantage that it provides a soft decision with interpretability. Fuzzy set describes the degree to which a pattern fits within a class, rather than a hard decision. Fuzzy systems are successfully used to handle many applications in the real world such as control systems and pattern classification problems [16]. In Fuzzy systems, the linguistic variables are used to create choices based on fuzzy rules and this is the reason why these systems get better results compared to those use crisp values ( value of 0 or 1).. Thus, fuzzy logic provides simple modelling, fast convergence, and noise reduction. There are two main types of fuzzy system which are A Mamdani-model and A Sugeno model. A Mamdani-Type fuzzy inference system is

implemented to predict a benign or malignant in this proposed method. Fuzzy inference systems consist of the following main parts;

### 1- Linguistic variables

In fuzzy system, each fuzzy label has membership function which assign membership degree. The numerical values and mathematical operations are converted into linguistic variables and logic operations. Fuzzy Operators are min and max. Each variable is converted to fuzzy set through fuzzification process.

### 2- Fuzzification

In this process, the fuzzy system translates a crisp of the input into fuzzy weights depending on fuzzy membership. The fuzzier is to determine the degree of membership of a crisp input into a fuzzy set as shown in Figure 3. The input is the features which are nine attributes of the dataset. The Membership Functions have multi types such as triangular, Gaussian, trapezoidal, 1-function and Gamma function, piecewise linear and singleton. The trapezoidal membership is used. Then, the features are converted into fuzzy sets by applying membership function with fuzzier. In the proposed fuzzy system, after applying the knowledge base, two output fuzzy sets are labeled “low” and "High". Each feature belongs to at least one fuzzy set (Low, High). No value lies outside the range of all sets. For each label, at least one element of the space should have all membership degrees equal to one. The intersection of two fuzzy sets “High” and “Low” has the membership function  $\mu_{High \cap Low}$ . Figure 2 shows that the black is for fuzzy output high and the red one is for the fuzzy output low.

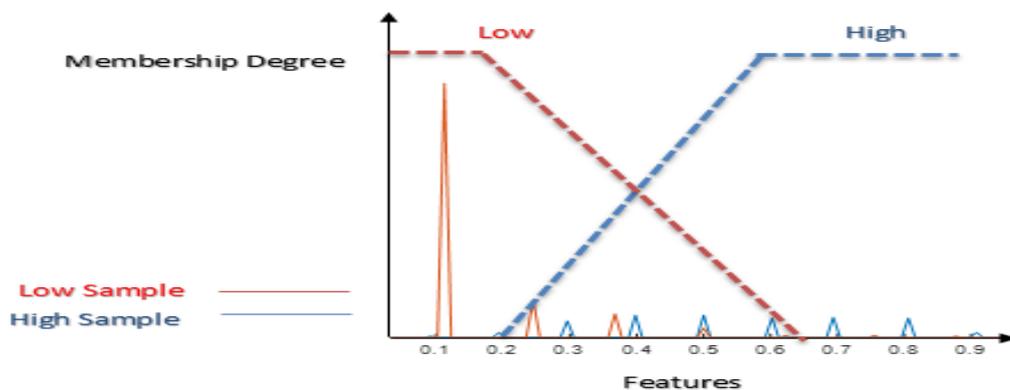


Figure 2. The fuzzification of the proposed fuzzy system

### 2- Knowledge base

Knowledge Base contains both linguistic inputs with their corresponding weights and a set of fuzzy rules. The decision making process is performed in the inference engine using the rule base. These fuzzy rules represent the relation between the input and output of the fuzzy variables. The objective of the knowledge base is to find the weights of the output. The form of the rule is [if antecedent then consequent]. The inference engine combines the weighted consequents of all relevant rules into a single fuzzy set using the aggregation operation. In the proposed method, there are only two linguistic outputs; benign and malignant.

- IF (Clump Thickness, Uniformity of Cell Size and Marginal Adhesion.....) are Low, Then diagnosis is Benign.

- IF (Clump Thickness, Uniformity of Cell Size and Marginal Adhesion..... are High, Then diagnosis is Malignant.

➤ Fuzzy logic Operations

Generally, the fuzzy logic operations are restricted by either intersection or union of fuzzy weight as shown the following equations (1), (2)

$$\mu_{\text{output-High}} = (\mu_{\text{High-Clump Thickness}} * \mu_{\text{High-Uniformity of Cell Size}} * \mu_{\text{High-Marginal Adhesion}}) \quad (1)$$

$$\mu_{\text{output-low}} = (\mu_{\text{low-Clump Thickness}} * \mu_{\text{low-Uniformity of Cell Size}} * \mu_{\text{low-Marginal Adhesion}}) \quad (2)$$

#### 4- Defuzzification

Defuzzifier is used to convert outputs of the fuzzy into crisp values. In The proposed FIS, the defuzzifier is shown following

- A. IF  $\mu_{\text{output-High}} > \mu_{\text{output-Low}}$  THEN output is Malignant.
- B. IF  $\mu_{\text{output-High}} < \mu_{\text{output-Low}}$  THEN output is Benign.

#### 3.1.3. Proposed Genetic Algorithm

Generally, a genetic algorithm is a stochastically iterative process that consists of a population of individuals which belong to evolutionary algorithms (EA). It is usually applied to spaces which are too large and complex to be exhaustively searched. Besides, it is used for many applications in bioinformatics and medical. It is used to generate solutions for optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover. In the space of problem, each one represents by a finite string of symbols, known as a chromosome. It represents a possible solution in a given problem space. Each position in the chromosome is called gene. An initial population of individuals is generated randomly or heuristically. In every evolutionary step (generation), the individuals in the current population are decoded and evaluated according to a fitness function. The fitness function describes the optimization problem in the search space. It is considered as basically a measure of the individual's ability to solve the problem. In the proposed GFS, fuzzy system classifier is improved by a genetic algorithm. The parameters for each feature of fuzzy membership are optimized such that the classification accuracy is as maximum as possible [17]. It results in higher accuracy compared to the fuzzy classifier since the parameters of the fuzzy membership could be optimized as shown in Figure 3. The initial population is generated with threshold values. Each chromosome consists of (2\*no of proposed features) genes which are "9" for high ( $\mu_H$ ), and "9" for low ( $\mu_L$ ).the limits have to be implemented to reduce the range of each feature to improve the accuracy.

A new population is created from the previous one through several steps [18]. Regeneration of population is repeated until an acceptable fitness has been reached or until a predefined number of generations have been exceeded. The fitness function of the proposed genetic algorithm is the classification accuracy that could be calculated by eq.3.

$$Fitness\ Function = \frac{NB + NM}{NB + NM + NFB + NFM} * 100 \quad (3)$$

Where NB: number of true benign.  
 NM: number of true malignant.  
 NFB: number of false benign.  
 NFM: number of false malignant.

### 3.2. Adaptive Neuro Fuzzy Inference System (Second Method)

Adaptive Neuro Fuzzy Inference System (ANFIS) was developed by Jang in 1993 which is a combination of neural network (NN) and (FIS). The neural network is used to determine the parameters of the fuzzy system. Moreover, ANFIS implements on FIS which is Sugeno model. Neural network is used to improve the parameter of fuzzy logic automatically by means of learning methods. Back propagation is learning algorithm, which is used to train the system. Fuzzy sets are encoded as (fuzzy) connection weights. The architecture and the learning procedure of ANFIS [12] [13] consist the nodes and the directional links that connect these nodes as shown in Figure 3. Some or all of these nodes could be made adaptive thus making the outputs, which depend on the parameters, belonging to these nodes. The learning rules define how these individuals is generated randomly or heuristically. The ANFIS structure has multi-layer as shown as in Figure (3). Layer 1 is The Input layer which is represent the features of WBCD. Layer2 is Fuzzification process. In this layer membership function (Gaussian) is applied to generate membership values for next layer. Layer 3 is represented as rule base which are generated by Takagi-Sugeno. Then Normalization is applied through layer 4. Layer 5 is the Defuzzification which the weighted values of each rule is calculated then summation layer in layer 6. ANFIS method is described as shown in Figure 4.

Firstly, in the proposed method Principle analysis component (PCA) is implemented as a feature selection.

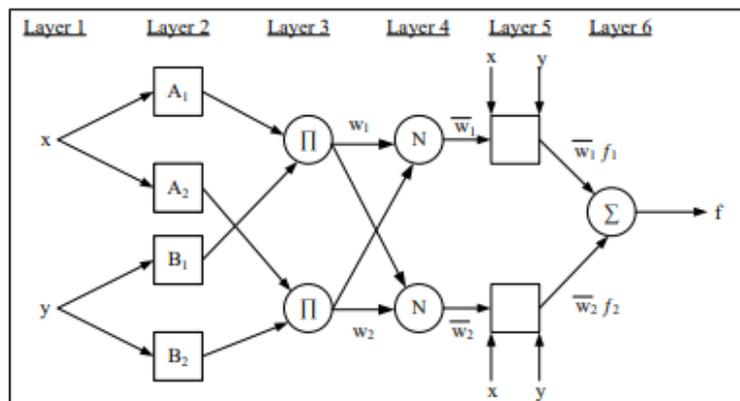


Figure 3. Adaptive-Neuro Fuzzy Inference System (structure of ANFIS)

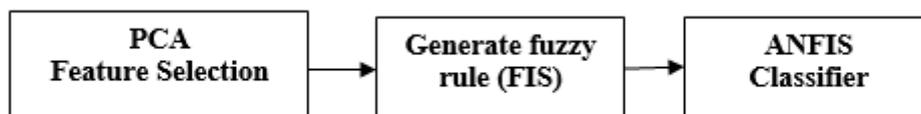


Figure 4. Adaptive-Neuro Fuzzy Inference System (Second Method)

- **PCA (Principle Component Analysis)**

Due to a large number of features which impact on the system, they make imbalanced datasets and this leads to the system to be unable to learn effectively [14]. This imbalance impacts on the performance of the classifier. So PCA will be implemented as it is statically technique which is used as a feature reduction process [9]. The process reduces the dimensionality of a dataset with minimal loss of information. Thus, feature reduction decreases the computation volume and this enhances the diagnosis accuracy.

#### 4. EXPERIMENTAL RESULTS

This section presents in details the simulation of the proposed of the two methods that implemented on the WBCD dataset using MATLAB R2016a.

##### 4.1. First method (Genetic fuzzy system)

###### A. Data Set

The WBCD will be divided into two sets.

- A. Training set contains all 699 cases of the database, and tests on the same testing set.
- B. Training set contains 75% of the data cases, and the testing set contains the remaining 25% of the cases

###### B. Classification

The features enter into inference engine that involves rules and membership function to obtain fuzzy output. Limits are implemented to each feature to improve accuracy as shown in Table 2.

Table 2. Limits of feature

Limit	v1	v2	v3	v4	v5	v6	v7	v8	v9
$\mu_L$	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
$\mu_H$	10	10	9	9	9	10	8	8	4

###### C. Implementation with GA

The parameters of the proposed GFIS are shown in Table 3. Also the constraints of proposed features are used to encode genes.

Table 3. GA parameters

GA parameters	Value of attributes
Genes number	n =20
Chromosomes	m=100
Crossover probability	cp =0.5
mutation Probability mp	mp =0.01
stopping crite ns when (a satisfaction fitness value)	ns =30

The proposed GA selects the best thresholds which gives the highest performance among all of the possible thresholds.GA achieves to “Th1” as shown in Table 4 which gives 97%.

Table 4. GFIS Result “Th1”

Th1	v1	v2	v3	v4	v5	v6	v7	v8	v9
ThL	0.76	1.97	1.36	0.66	1.89	0.85	0.87	0.98	0.88
ThH	8.54	4.92	6.12	6.90	4.53	7.37	7.19	3.87	3.05

#### D. implemented with Data Pre-processing (BLR)

Binary Logistic Regression is applied. Table 5 shows BLR output. As shown in Table 5, all coefficients are significant due to (Sig) is equal 0.0 and these means that p is less than .05. But v9 has less significant, because it has less score. Then Feature (v9) is ignored because it has less significant, the accuracy increases up to 97.7% with Th2 as shown in Table 6. GFIS results are shown in Table 7 with BLR and without BLR

Table 5. BLR OUTPUT

Feature v	Score	Df	Sig.
v1	358.348	1	.000
v2	467.608	1	.000
v3	468.786	1	.000
v4	339.386	1	.000
v5	325.870	1	.000
v6	471.600	1	.000
v7	400.155	1	.000
v8	354.596	1	.000
v9	125.172	1	.000
Overall Statistics	587.089	1	.000

Table 6: GFIS with BLR

Th2	v1	v2	v3	v4	v5	v6	v7	v8	v9
ThL	0.98	0.67	1.49	0.79	1.97	0.92	1.88	0.8	0.94
ThH	8.75	4.51	6.12	7.05	8.34	3.15	7.15	6.62	2.74

Table 7. GFIS Results

Th	Fitness	Accuracy
Th1 without BLR	0.03	97%
Th2 with BLR	0.023	97.7

## 4.2. Second Method (ANFS System)

### 4.2.1. PCA (Principle Component Analysis)

PCA is used as a feature reduction technique that depends on selecting the percentage of variation. It also depends on choosing uncorrelated with each other and captures as much original variance as possible. It is observed that the first three feature (Clump Thickness,

Uniformity of Cell Size, and Uniformity of Cell Shape) are most effected as shown in Table 8. They will be the input of the classification of the system of ANFIS.

Table 8. PCA Results

Attribute No.	Feature	Percentage of Importance
1	Clump Thickness	69.0859
2	Uniformity of Cell Size	7.1668
3	Uniformity of Cell Shape	6.0622
4	Marginal Adhesion	4.4344
5	Single Epithelial Cell Size	3.8973
6	Bare Nuclei	3.4375
7	Bland Chromatin	2.5304
8	Normal Nucleoli	2.2488
9	Mitoses	1.1367

#### 4.2.2. Classification

After PCA, The proposed system is generated using grid partitioning. The gradient descent method (backward pass) is used to adjust optimally the premise parameters. The resultant network has 27 rules, 3 inputs, 1 output, and 3 Gaussian membership functions per input as shown in Fig 5. The dataset is partitioned into 3 mutually exclusive sets.

- Training data consists of 200 cells.
- Checking data consists of 263 cells.
- Testing dataset consists of 200 cells.

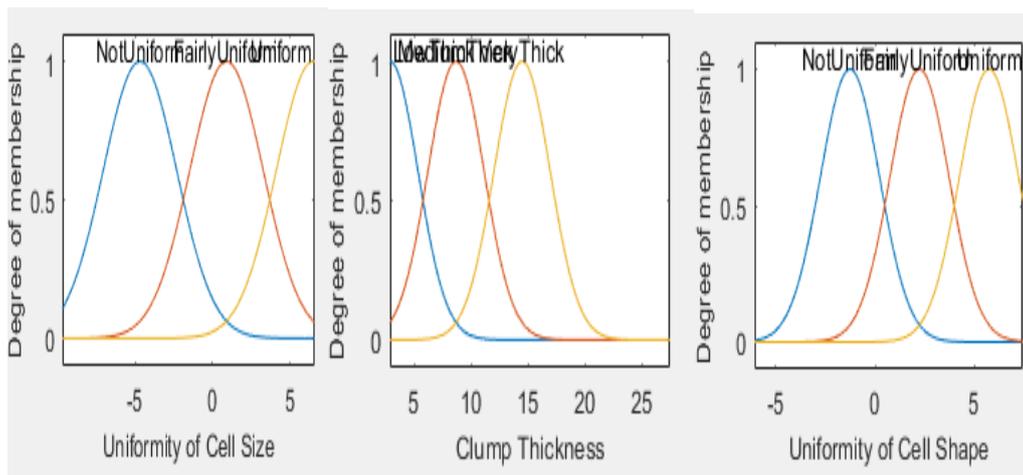


Figure 5. Membership functions of the generated fuzzy inference System

Table.9 shows the ANFIS system parameters. The back propagation learning rule is applied by ANFIS to train the network for classification. The network was re-trained with the same arameters. The results after the 300th epoch were promising. According to classification output,

there is no error in benign cases. It detects two cases error in malignant only as shown in Table 10. Figure. 6 shows the Results of re-training ANFIS with normalized data after 300 cycles with accuracy 99.1%

Table 9. ANFIS system parameters

ANFIS system parameters.	Value of Attributes
Number of nodes	78
Number of linear parameters	108
Number of nonlinear parameters	18
Total number of parameters	126
Number of training data pairs	201
Number of checking data pairs	285
Number of fuzzy rules	27

Table 10. Classification Error

Class	Correct	Incorrect
Benign	144	0
Malignant	56	2

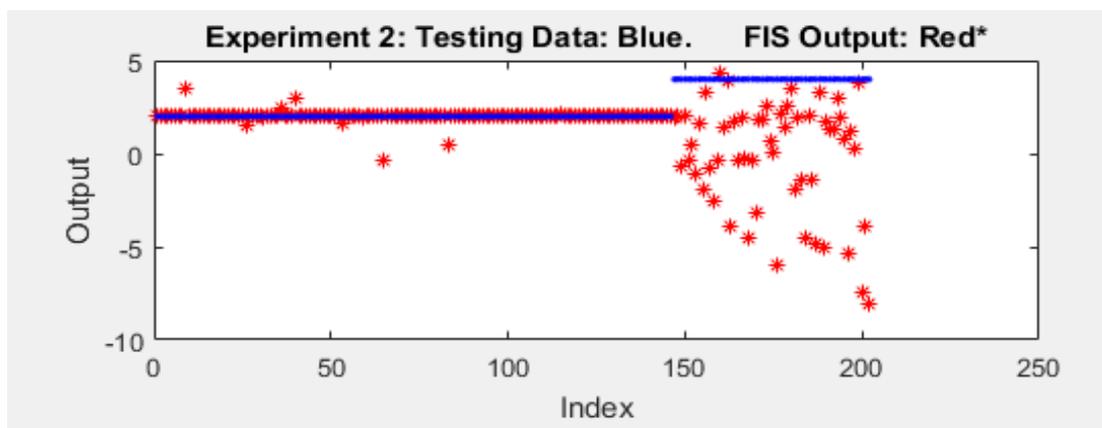


Figure 6. the Results of re-training ANFIS with normalized 3-dimensional data.

## 5. COMPARISON AND DISCUSSION

Proposed paper uses two classifier for diagnosing breast cancer which applied on WBCD as shown in Table 11. First classifier is GA-fuzzy logic inference system, in this method, GA evaluates the best thresholds which provides the highest accuracy with 97% and the accuracy is increased by implementing BLR up to 97.7% the second method is ANFIS that uses both ANN and fuzzy logic. ANFIS with feature selection optimizes the FIS parameters to get the highest accuracy with 99.1%. An ANFIS is better than GFIS because it provides the mapping relationship between the input and output information using hybrid learning strategy to discover out the optimal distribution of membership functions. Moreover, ANFIS model has a hybrid

learning algorithm that consists of both least squares method and back propagation learning which provide a trained model with ANN that is used to optimize the parameters of a fuzzy model. ANFIS makes sufficient rules concerning input and output information Moreover, ANFIS consists a multi-layer so it provides the best accuracy which is 99.1%.

Table 11. GFIS, ANFIS Results

No	Method	Accuracy
1	Genetic-fuzzy inference system (GFIS)	97.7
2	Adaptive-neural fuzzy inference system(ANFIS)	99.1

## 6. CONCLUSIONS

The proposed system implements two classifier on breast cancer diagnosis using the Wisconsin database (WBCD). First classifier is the Genetic-fuzzy inference system (GFIS) with applying Binary Logistic Regression. Classification accuracy obtained for the Genetic-fuzzy classifier is 97.7%. However, the second classifier of Adaptive Neuro Fuzzy inference system (ANFIS) has higher accuracy than GFIS. ANFIS consists of multi layers but Genetic-fuzzy classifier does not consist multi-layer. Also, the selection of a suitable feature reduction method (PCA) optimizes ANFIS classifier. ANFIS achieves up to 99.1%, because ANFIS makes sufficient rules concerning input and output information. Also, ANFIS consists of three layers input layer, hidden and output layer using hybrid learning strategy to discover out the optimal distribution of membership function

In future work, a complex medical diagnosis tool will be built for multi-disease such as Autism diagnosis, lung cancer diagnosis, lever cancer diagnosis.....and brain cancer diagnosis using multi classifier such as GFIS,ANFIS and hybrid genetic-neuro fuzzy inference system to achieve the possible highest accuracy with high confidence.

## REFERENCE

- [1] Torre, L.A& Siegel, R.L.; Ward, E.M.; Jemal(2016), A. Global Cancer Incidence and Mortality Rates and Trends—An Update.Cancer J.Clin.
- [2] Fatima S. Ahadi1& Mark R. Desai1, Chengwei Lei (2017), Feature-Based Classification and Diagnosis of BreastCancer Using Fuzzy Inference System © IEEE
- [3] Abir Alharbi &F. Tchier (2017), Mathematical Biosciences, Using a Genetic-Fuzzy Algorithm as a ComputerAided Diagnosis Tool on Saudi Arabian Breast Cancer Database.
- [4] Nilashi, M& Ibrahim, O., Ahmadi, H., Shahmoradi (2017) Telematics and Informatics, A Knowledge-Based System for Breast Cancer Classification Using Fuzzy Logic Method.
- [5] Nguyen,Khosrav & Creighton, Nahavandi, Medi-cal data classification using interval type-2 fuzzy logic system and wavelets.
- [6] Hossein Ghayoumi & Zadeh,Omid Pakdelazar (2012) Middle East Journal of Cancer ,Diagnosing Breast Cancer with the Aid of Fuzzy Logic Based on Data Mining of a Genetic Algorithm in Infrared Images
- [7] Payam ZARBAKHSH & ZABDOJLIL, Addeh,Hasan Demirel(2017) Early detection of breast cancer using optimized ANFIS and features selection © IEEE

- [8] A. Onan (2015) A fuzzy-rough nearest neighbor classifier combined with consistency-based subset evaluation and instance selection for automated diagnosis of breast cancer© Elsevier.
- [9] O.Inan & M. S. Uzer, N. Yılmaz (2013) A new hybrid feature selection method based on association rules and pca for detection of breast cancer @ International Journal of Innovative Computing, Information and Control.
- [10] SudiptaRoy & Shayak Sadhu(2016)" Brain Tumor Classification using Adaptive Neuro-Fuzzy Inference System from MRI" 2016 SERSC
- [11] K. Bache, & M. Lichman, UCI machine learning repository. URL(2013) <http://archive.ics.uci.edu/ml/machine-learning-databases/breastcancer-wisconsin>.
- [12] M. Ashraf (2010) Information Gain and Adaptive Neuro-Fuzzy Inference System for Breast Cancer Diagnoses@IEEE
- [13] Ahmad Taher Azar(2014) Dimensionality reduction of medical big data using neural-fuzzy classifier@ Springer.
- [14] Gouda I. Salama (2012) Experimental Comparison of Classifiers for Breast Cancer Diagnosis@ IEEE
- [15] [.https://www.wcrf.org/dietandcancer/cancer-trends/breast-cancer-statistics](https://www.wcrf.org/dietandcancer/cancer-trends/breast-cancer-statistics)
- [16] Salma Elhag & Alberto Fernández, Abdullah Bawakid, Saleh Alshomrani, Francisco Herrera (2015 Z) On the combination of genetic fuzzy systems and pairwise learning for improving detection rates on Intrusion Detection Systems@ ELSEVIER
- [17] Mrs. Meenakshi M. Pawar & Dr. Sanjay N. Talbar (2016) Genetic Fuzzy System (GFS) based Wavelet Co-occurrence Feature selection in Mammogram Classification for Breast Cancer Diagnosis.
- [18] M.E. Cintra a & H.A. Camargo, M.C. Monard (2016) Genetic generation of fuzzy systems with rule extraction using formal concept analysis@ELSEVIER.
- [19] Hamza Turabieh & Muhanna Muhanna(2015) GA-based feature selection with ANFIS approach to breast cancer recurrence@ International work-conference on Bioinformatics and biomedical engineering.
- [20] M. Hanefi CALP (2019) A Hybrid ANFIS-GA Approach for Estimation of Regional Rainfall Amount@GU J Sci, 32(1): 145-162