# Breast Cancer Classification Using Adaptive Neuro Fuzzy Inference System (anfis)

Omoju A. Thomas
oathomas@memphis.edu
Department of Electrical and Computer Engineering
The University of Memphis, Memphis, TN USA 38152

Abstract – **In this paper an application of the Adaptive Neuro Fuzzy Inference System for breast cancer classification is presented.**

## I. INTRODUCTION

Cancer is the second cause of all deaths in the United States. Among the cancers in women, breast cancer accounts for 40, 900 deaths in the year 2001 alone [3]. Furthermore, breast cancer is the second deadliest cancer in women. Early detection is the key to defeating the disease. A technique called fine needle aspiration has been developed to allow for the early detection of cancerous cells in the body. Assuming that there exist a repository of data that represents cases for both victims and non-victims alike, then the research community can use a pseudo-statistical method to classify and determine whether a biopsy contains a cancerous cell or not. Adaptive neuro fuzzy inference system is a computational tool that is capable of approximating a function that solves the problem of breast cancer classification.

## II. FOUNDATIONS OF ANFIS

The anfis network is a superset of all feed forward neural networks using either a pure back propagation gradient-descent learning rule, or a hybrid learning rule that uses back propagation and a least squares method [1]. Anfis allows for the membership functions of the fuzzy inference system generated to be learnt by applying its learning algorithm. This mechanism is what sets the anfis network apart from standard fuzzy inference systems. Anfis generates a Sugeno-type fuzzy system. An in depth treatment of the Sugeno-type fuzzy system is beyond the scope of this paper.

## III. PREPROCESSING OF DATA

The data that was used to run the experiments that are presented in this paper were retrieved form the University of California, Irvine machine learning repository. The data referred to as the Wisconsin breast cancer data was collected over a series of four years by Dr. William H. Wolberg of the University of Wisconsin Hospitals, Madison [2]. The data consist of 10 attributes as displayed in table 1. The data represents 699 instances of which 65.5% is benign and 34.5% is malignant. The attributes represent the clump thickness, the uniformity of the cell and so on as displayed in table 1. The value of 2 has been assigned as the class label for all instances that are benign, and the value of 4 has been assigned as the class label for all instances that are malignant.

A large percentage of the accuracy of any method performed on any data lies within the data itself. If data overlaps, or data has other undesirable qualities, these degrades the performance of algorithms that utilizes the data. As a result, it is best to first pre-process the data, before an algorithm is run on it. Principal components analysis (PCA) using the covariance matrix was performed on the breast cancer data to see which dimensions contained the most information. As illustrated in table 1, the first five attributes of the data contain 90.5 percent of the information. The dataset contained some missing attributes; as a result, those rows of data that had missing attributes were deleted reducing the dataset to 683 instances.

TABLE I
PCA Results

| Percentage of Importance | Feature |
|---|---|
| 69.0859 | Clump Thickness |
| 7.1668 | Uniformity of Cell Size |
| 6.0622 | Uniformity of Cell Shape |
| 4.4344 | Marginal Adhesion |
| 3.8973 | Single Epithelial Cell Size |
| 3.4375 | Bare Nuclei |
| 2.5304 | Bland Chromatin |
| 2.2488 | Normal Nucleoli |
| 1.1367 | Mitoses |

Figure 1 shows the distribution of the data after clipping the dimensions down to five.

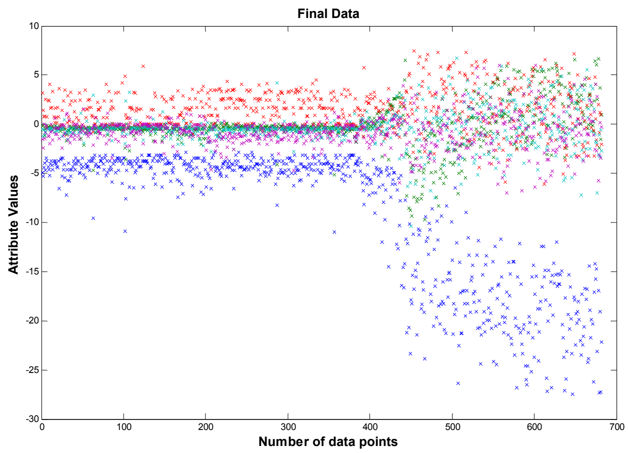


Figure 1. The distribution of Wisconsin dataset after principal component analysis.

The first five attributes were selected as the inputs to an adaptive neuro fuzzy inference system.

## IV. EXPERIMENTS

A series of 5 experiments were carried out on the dataset to approximate an adequate condition for anfis classification of breast cancer.

*A.* An anfis network was generated using grid partitioning. The resultant network had 234 rules, 5 inputs, 1 output, and 3 Gaussian membership functions per input. The dataset was partitioned into 3 mutually exclusive sets. The first set contained the training data which consisted of 146 instances of benign cells and 82 instances of malignant cells. The second set contained the checking data that is used by anfis. It consisted of 147 instances of benign cells and 73 instances of malignant cells. The final set contained the testing dataset which likewise consisted of 147 instances of benign cells and 82 instances of malignant cells. Figure 2 shows the membership function generated as a result of the network before training.

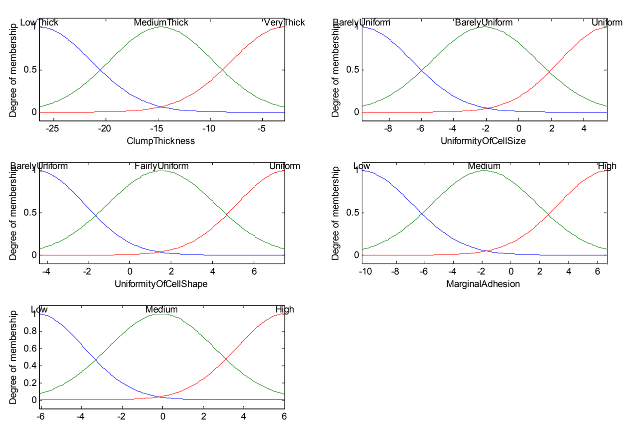


Figure 2. Membership functions of the generated fuzzy inference system.

The back propagation learning rule was applied by anfis to train the network for classification. Figure 3 shows the results of classification, and table 2 shows the classification distribution.

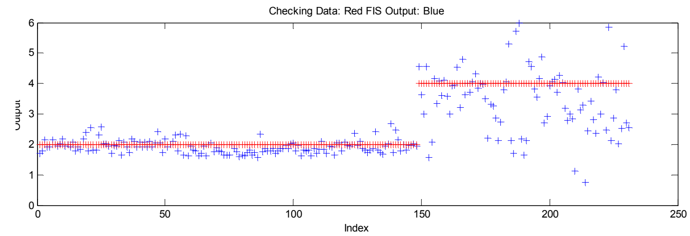


Figure 3. Results of training anfis with the back propagation learning rule after 300 epochs.

TABLE II
Experiment 1: Classification Error

| Classes | Correct | Incorrect |
|---|---|---|
| Benign | 145 | 1 |
| Malignant | 71 | 12 |

*B.* For the second experiment, the 5-dimensional dataset was normalized using the z-score algorithm. The network was re-trained with the same parameters as before those are three Gaussian membership functions per input, and a training cycle of 300 epochs. The results after the 300$^{th}$ epoch was promising, resulting in a classification error of 3.46%. The network was allowed to continue training for another 300 epochs, and the results improved to achieve a classification error of 0.87%. Figure 4 displays the results of classification after 300 cycles as well as after 600 cycles. Table 3 show the classification distribution of the experiment.


Figure 4. Results of re-training anfis with normalized data after 300 cycles and 600 cycles respectively.

TABLE III
Experiment 2: Classification Error

| Class | Correct | Incorrect |
|---|---|---|
| Benign | 147 | 1 |
| Malignant | 76 | 7 |
| After 6oo Epochs | | |
| Benign | 146 | 2 |
| Malignant | 83 | 1 |

*C.* For the third experiment, the 5-dimensional normalized dataset was further reduced to 3-dimension since the first 3 dimensions of the data contain 82% of the data. The reason for this experiment was to discover whether the other dimensions that contained minimal information could be considered as noise. Figure 5 shows the distribution of the dataset after the reduction.

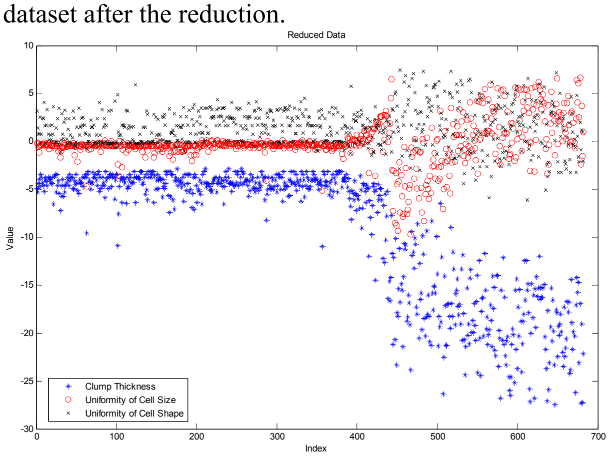


Figure 5. The distribution of 3 dimensional normalized dataset

An anfis network was generated using grid partitioning. The resultant network had 27 rules, 3 inputs, 1 output, and 3 Gaussian membership functions per input. All other variables were as before in experiments 1 and 2. Figure 6 shows the network's structure.

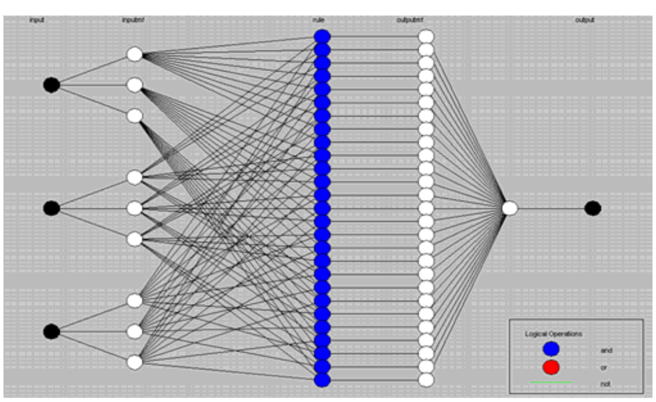

Figure 6. Network structure of ANFIS network with 3 inputs, 1 output and 27 rules.

After training of the network using the back propagation rule for 600 epochs, the result of the experiment did not improve the classification error for the dataset; however, this approach degraded the classification error to 3.47%. Table 4 shows the classification distribution, and figure 4 shows the results of the classification.

TABLE IV
Experiment 3: Classification Error

| Class | Correct | Incorrect |
|---|---|---|
| Benign | 141 | 7 |
| Malignant | 82 | 1 |

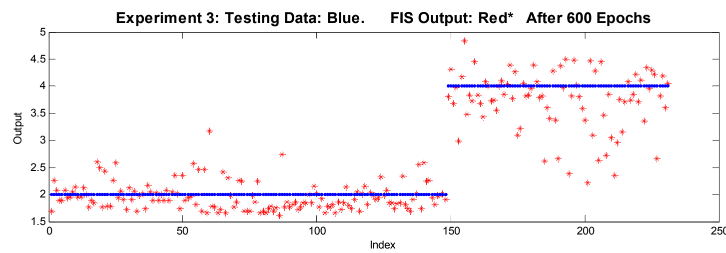


Figure 6. Results of re-training anfis with normalized 3-dimensional data.

*D.* For the fourth experiment, the anfis network of the third experiment was re-trained using the anfis hybrid learning rule which combines back propagation and a least-squares method. This experiment did not decrease the classification error; instead it degraded the solution, increasing the classification error to 8.658%. Figure 7 shows the adapted membership functions, figure 8 shows the classification results, while table 5 shows the distribution.

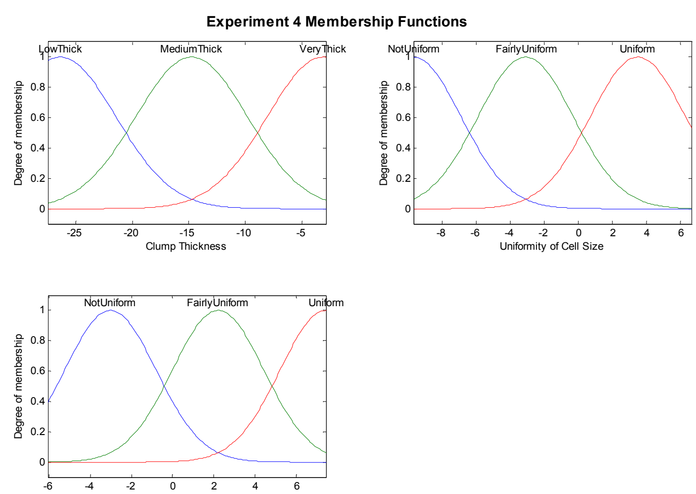


Figure 7. Adapted Membership functions after training

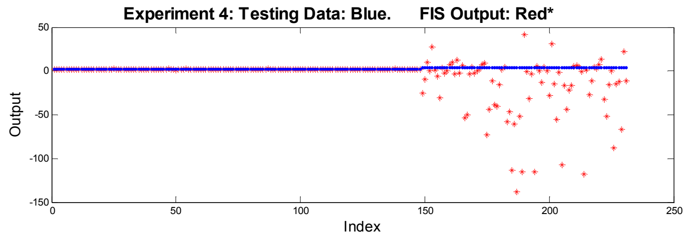

Figure 8. Results of re-training anfis with normalized 3-dimensional data using the hybrid learning rule.

TABLE V
Experiment 4: Classification Error

| Class | Correct | Incorrect |
|---|---|---|
| Benign | 146 | 2 |
| Malignant | 65 | 19 |

## V. CONCLUSIONS

In this paper, an application of the adaptive neuro fuzzy inference systems was applied to the problem of breast cancer classification of a 9 attribute dataset. Table 6 summarizes the conclusions of the 4 experiments that were carried out in search of a solution.

TABLE VI
Summary of experiments

| Exp. | Algorithm | Class. error | dimension | Epoch |
|---|---|---|---|---|
| 1 | Back prop | 6.90 % | 5 | 300 |
| 2 | Back prop | 3.4632 % | 5 | 300 |
| 2 | Back prop | 0.8658 % | 5 | 600 |
| 3 | Back prop | 3.4632 % | 3 | 600 |
| 4 | Hybrid | 8.658 % | 3 | 10 |

The conclusion was met that a reduction in the dimension of the data from 9 to 5 was desirable in attaining a good classification for benign and malignant cancer cells based on the Wisconsin dataset.

Furthermore, normalizing the clipped data applying a generalized Z-score algorithm reduced the classification error of a 5 dimensional anfis network from 6.9% to 0.9%.

Re-clipping the data from 5 dimensions to 3 dimensions diminished the performance of the network there by leading to the conclusion that the marginal adhesion and single epithelial cell size are important factors in classifying breast cancer in humans.

Finally, the following five attributes are important in determining the presence or absence of breast cancer: clump thickness, uniformity of cell size, uniformity of cell shape, marginal adhesion and single epithelial cell size.

## REFERENCES

[1] J.S. Jang, "*Anfis: Adaptive-network-based fuzzy inference system*," IEEE Trans. Syst., Man, Cybern., vol. 23, pp. 665--685, Mar. 1993

[2] W. H. Wolberg and O.L. Mangasarian: "*Multisurface method of pattern separation for medical diagnosis applied to breast cytology*", Proceedings of the National Academy of Sciences, U.S.A., Volume 87, December 1990, pp 9193-9196.

[3] American Cancer Society [online] www.cancer.org