

# **Breast Cancer Detection using Neuro-Fuzzy**

**Soft Computing Project Report**

*Submitted by -*

**ANANDMOHAN TRIPATHI**

*Registration No.-11812654*

*In partial fulfillment of Community Development Project for the award of the  
degree of*

**“BACHELOR OF TECHNOLOGY”**

**School of Computer Science and Engineering**

**LOVELY PROFESSIONAL UNIVERSITY**

**Phagwara, Punjab**

# INTRODUCTION

Breast cancer is common disease which can cured with early detection. There comes role of technology. We can predict whether cell is dangerous (malignant) or not dangerous (benign). Neuro-Fuzzy classification is a type of Hybrid system which is combination of fuzzy and neuro system

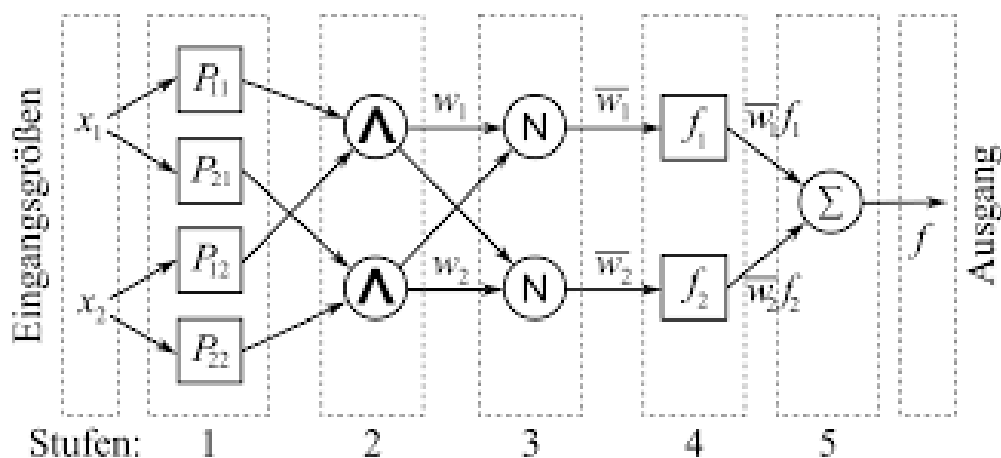
Due to the presence of ambiguity in the training dataset, overlying borders among classes and vagueness in describing characteristics some uncertainties can still occur at any stage of a classification system. The fuzzy set theory approach is flexible enough to manage the various facets of uncertainties about real life situations. ANN and fuzzy set theory based combined approach is the neuro-fuzzy technique. A typical neuro-fuzzy classification based approach exploits these two techniques in an efficient way. It combines the human like rational thinking of fuzzy based systems with the connective structure and the learning ability of ANNs to develop a Neuro-Fuzzy System (NFS).

## Methodology

- Step1: Collecting the data which is available uci (breast cancer wisconsin data)
- Step2: Analysis of Dataset as it involves many correlation and removing unwanted columns which are not necessary as less the unnecessary feature more accurate the model
- Step3: Use pair plot and heatmap to check relation and affect for features in diagnosis
- step4: Once the data is cleaning and data is structured and ready for fit it in Neuro-Fuzzy model. Anfis is used for Neuro-Fuzzy.
- Step5: Train and test and check Accuracy and error for model

# Neuro-Fuzzy

In this we use methodology of Fuzzy using Neuron. It consists of Five Layers the layers from 1 to 5 are called: input linguistic layer, condition layer, rule layer, consequent layer, output linguistic layer. The fuzzification of the inputs and the defuzzification of the outputs are respectively performed by the input linguistic and output linguistic layers while the fuzzy inference is collectively performed by the rule, condition and consequence layers



## DATASET ANALYSIS

The dataset consists of 32 rows the columns diagnosis classifies the data in two parts malignant and benign. Apart from it consists of other 31 columns such as radius\_mean, smoothness, fractal, concavity, compactness etc. With help of matplotlib library we can clean data and remove correlation between features by finding relation between features. After it we plot these points and find out grouping for member function

- ID number.
- Diagnosis (M = Malignant, B = Benign).
- Exactly 10 real-valued characteristics are calculated for each of the cell nucleus as:
  - The radius (the mean value of distances from the center to points on the perimeter).
  - The texture (the standard deviation of gray-scale values).
  - The perimeter.
  - The area.
  - The smoothness (the local variation in the radius lengths).

- The compactness ( $[\text{perimeter}]^2/\text{area} - 1.0$ ).
- The concavity (the severity of concave portions of the contour).
- The concave points (the number of concave portions of the contour).
- The symmetry.
- The fractal dimension (coastline approximation - 1).

On plotting graph matrix will show the relation between features. From this we can identify which feature to keep and which to discard

By using heatmap we can get to see correlation

After deleting the unwanted and correlated column we plot feature with respect to diagnosis for Membership function.

Once the membership function are done we just need to fit the data and check for error and accuracy

We need to split our data into two part one to train and other to test as it is important for model to work fine both training as well as test data

## Library Used

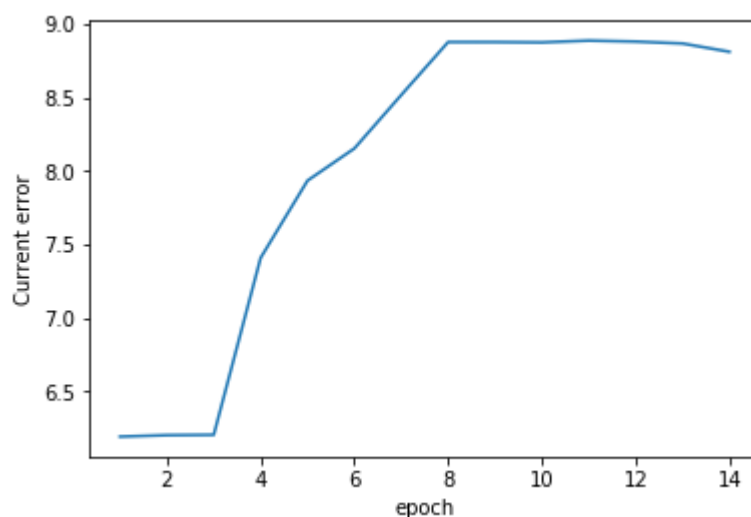
- Numpy:- representing feature
- Pandas:- data frame
- Matplotlib:- for pair plot
- Seaborn:- for heat map and relation
- ANFIS :- fitting data and checking error

# Graph

## Error

```
current error: 6.190382586490118
current error: 6.2003280995089485
current error: 6.202547068547006
current error: 7.407043120115204
current error: 7.932903836770583
current error: 8.152054691097067
current error: 8.515664569332255
current error: 8.876056987841888
current error: 8.875822887109338
current error: 8.873675728545804
current error: 8.885846624801605
current error: 8.879155985926737
current error: 8.866191736928819
```

## Error graph



**Github link:-** <https://github.com/anandtripathi10/breast-cancer-detection-using-Anfis/tree/main>