12/11/2017

Abdul Rahman Khan

Neural networks and fuzzy system

Breast Cancer Classification Using Neural Networks

Abstract:

Breast cancer diagnosis posture a great challenge to the researchers. Data mining and machine learning have been used to solve many real world complicates problems. Breast cancer diagnoses discriminates between malignant and benign cancer cases. Breast cancer diagnoses is a classification problem. In this paper I assessed the use of artificial neural networks to solve breast cancer diagnosis problem.

Introduction:

Breast cancer is growing at an alarming rate. In US 1 out of 8 women’s have a risk of developing breast cancer in their lives [1] Breast cancer is a complicated disease to diagnose and rectify. It can be cured if detected on early stage. So we need an accurate and robust process to detect breast cancer.

Artificial Neural networks:

Artificial neural networks is a powerful machine leaning technique used to solve many real world problems.

In computer world everything is considered to be rule base. But this technique is changing the trend. It can give you a dynamic solution to your problem. Artificial neural networks (ANN’s) has also been widely used to solve computation extensive and classification problems.

Artificial Neural Network topologies: [2]

Feed-Forward ANN:

Feed forward network is a unidirectional artificial neural network. One neuron sends information to other neuron and it does not receive back any information.

* No feedback loops.
* Fixed inputs outputs.
* Used in pattern recognition/generation/classification.

Feed-Back ANN:

Feedback is a bidirectional ANN. Neurons can send information back to neurons in previous layer.

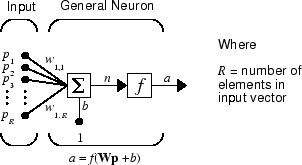
* Loops are allowed.

Multilayer Neural Network Architecture: [3]

Multilayer neural network is a feedforward network consists of neurons.

Neuron Model:

It follows the brain neuron model to simulate the process. It can take one or more inputs and sums them to produce an output. It can use any differential transfer function to conjure output.



Multilayer networks often use log-sigmoid (logsig), tan-sigmoid (tansig) and linear (pure-line) transfer functions.

Log-sigmoid and tan-sigmoid are used to solve pattern recognition problems. While pure-line is used to solve linearly separable problems

*Log-sigmoid:*

Input: Positive infinity to negative infinity.

Output: 0 to 1

*Tan-sigmoid:*

Input: Positive infinity to negative infinity.

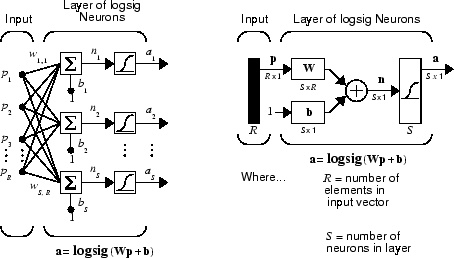
Output: 1 to -1

*Linear transfer function:*

Input and output ranges in positive to negative infinity.

Feedforward Neural Network:

Feedforward neural network consists (possibly) of multiple neurons.



Data Selection:

Wisconsin Breast Data Set: [4]

Wisconsin Breast Cancer Database (WBCD) contains the data of various breast cancer cases. In this data, cases are categorized in two types (malignant and benign).

Data consists of 11 attributes. First attribute contains the id of patient.

From attribute 2 to 10 contains different properties for detection of breast cancer ranging values from 1 to 10.

|  |  |  |
| --- | --- | --- |
| No | Attribute | Domain |
| 1 | Clump thickness | 1 – 10 |
| 2 | Uniformity of cell size | 1 – 10 |
| 3 | Uniformity of cell shape | 1 – 10 |
| 4 | Marginal adhesion | 1 – 10 |
| 5 | Single epithelial cell | 1 – 10 |
| 6 | Bare nuclei | 1 – 10 |
| 7 | Bland chromatin | 1 – 10 |
| 8 | Normal nucleoli | 1 – 10 |
| 9 | Mitosis | 1 – 10 |
|  | Class | 2 for benign, 4 for malignant |

Missing Values:

Wisconsin data set does contains the samples of 699 cases breast cancer cases. Some of these samples have missing values against bare nuclei. I can treat these samples in three different ways.

Drop:

I can remove these samples from data set. It may not even effect the data but not a good approach though.

Neglect:

Use all data set (contain affected samples) as testing data. This may effect on our experiments outcome.

Replace with median:

Replacing missing values with the median of that attribute is somehow more suitable. Because in this way we will be using samples data (not completely discarding samples).

Bare nuclei contains 16 missing value. So I will take the median of bare nuclei and replace missing values with it.

Now I can use Wisconsin data set to find a resilient solution for the detection of breast cancer.

HYPOTHESIS:

1: Changing no of layers will affect the performance of neural network.

We have to classify the cancer in two forms which are malignant and benign. So I think that our problem can be solved with three layered neural network and it will give us the best result. As it can easily cover meshed regions.

Parameters and Values:

|  |  |
| --- | --- |
| Parameters | Values |
| Functions | Tansig, trainr, learngd, mse |
| Goal | 0.01 |
| Epochs | 100 |
| Learning rate | 0.001 |
| Max-fail | 200 |

Experiments:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Hidden Layers | Data Type | Training Data | Testing Data | MSE |
| 1 | Unformatted | 1-400 | 401-699 | 4.68% |
| 3 | Unformatted | 1-400 | 401-699 | 1.34% |
| 6 | Unformatted | 1-400 | 401-699 | 1.0075% |
| 9 | Unformatted | 1-400 | 401-699 | 1.5100% |
| 12 | Unformatted | 1-400 | 401-699 | 2.5150% |

Conclusion:

Increase in hidden layers results in decrease in MSE. But not from a certain threshold. It is because of over-fitting problem. Over fitting occurs when you try to stimulate a problem with a complex solution instead of simple optimal solution.

As clear from the experiments that no. of layers does effect the performance but it can vary with the problem.

2: Increasing the training data will decrease MSE.

If I increase the training dataset MSE will decrease. Because network will get more samples to train. And it will be able to distinguish between different variety of cases samples.

Parameters and Values:

|  |  |
| --- | --- |
| Parameters | Values |
| Functions | Tansig, trainr, learngd, mse |
| Goal | 0.01 |
| Epochs | 100 |
| Learning rate | 0.001 |
| Max-fail | 200 |

Experiments:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Hidden Layers | Data Type | Training Data | Testing Data | MSE |
| 3 | Unformatted | 1-200 | 201-699 | 4.68% |
| 3 | Unformatted | 1-300 | 201-699 | 1.34% |
| 3 | Unformatted | 1-500 | 501-699 | 3.011% |
| 3 | Unformatted | 1-600 | 601-699 | 2.0087% |

Conclusion:

Increase in training data does improves the performance of our neural network. But sometimes it gives us random results. So I can’t be certain about my hypothesis because of non-linear output.

This paper provides the insight of experiments performed on Wisconsin data set using artificial neural network. From the above experiments it is observed that we can use ANN for better accuracy and performance breast cancer diagnosis.

Experimental Detail:

Used MATLAB for conducting experiments. MATLAB provides nntool for such experiments. It is quit handy for researchers to perform rigorous experiments.

# References

|  |  |
| --- | --- |
| [1] | "US Breast Cancer Statistics," 3 12 2017. [Online]. Available: http://www.breastcancer.org/symptoms/understand\_bc/statistics. |
| [2] | "Artificial Intelligence - Neural Networks," Tutorials Point, 12 12 2017. [Online]. Available: https://www.tutorialspoint.com/artificial\_intelligence/artificial\_intelligence\_neural\_networks.htm. [Accessed 12 12 2017]. |
| [3] | "Mathworks," 3 12 2017. [Online]. Available: https://www.mathworks.com/help/nnet/ug/multilayer-neural-network-architecture.html. |
| [4] | "Breast Cancer Wisconsin (Original) Data Set," UCI, 12 12 2017. [Online]. Available: http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Original). [Accessed 12 12 2017]. |
| [5] | "Wikipedia," 3 12 2017. [Online]. Available: https://en.wikipedia.org/wiki/Artificial\_neuron. |