Breast Cancer Classification.

Name: Muhammad Arif
Year 4.
Supervisor: Junaid Akhter

Table of Contents

There is a space in the start, remove it.

Instead of left alignment, use justified.

Introduction

Breast cancer is always the dangerous disease for women in every country and unfortunately it's increasing day by day in previous years. Keeping this fact in mind many scientist give different effortless solution to overcome this problems. In the same way some researcher proposed the solution of Neural Network as a solution for this problem.

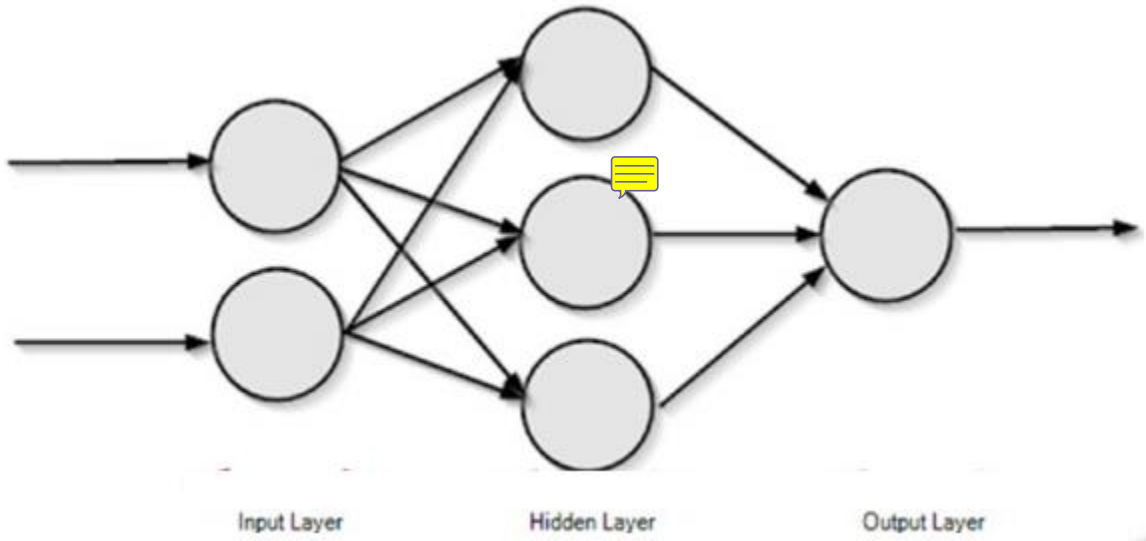
In this report I will explain how the neural network will work for the problem of breast cancer, how we solve problem on the basis of some hypothesis, experiment, analysis and then show the result.

Background

Neural network is copy design of human brain. It's work like human brain. Most important part of this network is Neuron which is called hidden layer of network. There is also two other Input and Output Layer.

Neural Network

Following figure is for neural network.



[2]

Preprocessing

Collect data from Wisconsin Diagnostic Breast Cancer (WBDC)[1] dataset and save it into cw.m file. Just put square bracket around the data and assign to new variable data. Now replace all missing element from 0. And then separate data into input and output matrixes just ignoring the ID column of all rows.

Network Architecture

We used backpropagation Network for this for network along with Neff function. We used different function like transig, learngd, trainrp and logsig as activation functions. Then we set different require attributes Goal is 0.01, learning Rate is 0.1 and number of epochs for running is 1000.

Post processing

After running the complete network we have to find percentage of accuracy and error for analysis of our solution so we find accuracy and error % by using:

1) Accuracy percentage = (Number of Element – Number of Error)/Number of Element *100
2) Error percentage = (Number of Error)/Number of Element *100

Experimental Result and Analysis

Experiment 1

In first experiment we will change the percentage of data for training and check the output.

Hypothesis

Initially hypothesis for this experiment is when training data is maximum accuracy will also be if training data will reduce then number of error will also increase and accuracy will decrease.

Analysis

Data distribution and accuracy and error.

Training Data	Test Data	Accuracy %	Error %
1-100	101-699	95.42	4.17
1-200	201-699	94.99	5.01
1-300	301-699	95.73	4.26
1-400	401-699	96.32	3.67
1-500	501-699	98.49	1.50

Result

Hypothesis for this experiment is approximately true because when we increase data for testing our out for accuracy is increasing and number of error is decreasing.

Experiment 2

In this experiment we will change the value of learning rate and check the output.

Hypothesis

Initially hypothesis for this experiment is when learning rate is high accuracy will be low and when learning rate will decrease then accuracy will increasing.

Analysis

Learning Rate and accuracy and error.

Learning Rate	Accuracy %	Error %
0.1	97.24	2.75
0.2	96.99	3.01
0.3	96.74	3.26
0.4	96.98	3.02
0.5	96.49	3.51

Result

Hypothesis for this experiment is approximately true because when we increase Learning Rate for testing our output for accuracy is increasing and number of error is decreasing.

Experiment 3

In this experiment we will change the number of Neuron of hidden layer and will check the out for different number of neuron.

Hypothesis

Initially hypothesis for this experiment is when we increase the number of neuron for hidden layer our neural network will become more complex but it could perform more accurate result as compared to less number of neuron.

Analysis

No. of Neuron, accuracy and error.

No. of Neuron	Accuracy %	Error %
5	94.23	5.76
10	96.24	3.07
15	96.74	3.26
20	96.99	3.00
25	96.74	3.25
30	95.24	4.76

Result

Hypothesis for this experiment is approximately true as we will use need base number of neuron because when we use more and more neuron in hidden layer our network become more complex. At some stage it will be over lifting the data and then we cannot get our desire output. Like when we use 50 and 100 neuron then our output is decreasing.

Experiment 4

In this experiment we will run our program different time in same input and will check the output.

Hypothesis

Initially hypothesis for this experiment is when run our program many times in same input our result should be same.

Analysis

No. of Neuron, accuracy and error.

No. of Neuron	Accuracy %	Error %
1	94.74	5.26
2	96.25	3.75
3	96.75	3.25
4	96.99	3.00
5	96.49	3.50

Result

As we have seen our output is different for every time we run our program even our data and network is same. It's happen because in every time our network get random number of weight. That's why our hypothesis is not compatible with result. To solve this problem we can save the random weight and then use it for our program.

Conclusion

Neural Network is already use for the solution of Breast Cancer. Main purpose of this activity to find some better solution for this problem. So we solve this by changing data for testing and training, changing number of neuron for hidden layer, changing learning rate and finally to find error in random weight assigning to network.

References

1) http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Original)
2) https://www.google.com.pk/url?sa=i&rct=j&q=&esrc=s&source=images&cd=&ved=0ahUKEwi12PW7tu7XAhVp7oMKHdq_BDUQjRwIBw&url=https%3A%2F%2Fdzone.com%2Farticles%2Ffeed-forward-neural-network-with-mxnetr&psig=AOvVaw3j2y1Zrbk2foeaHqIE9LCW&ust=1512400311954041