CW_BREAST CANCER

NNFS

NOVEMBER 27, 2017

MUHAMMAD ALI ABID RAZA
14031244

Table of Contents

Abstract

The ability in machine to learn itself and then show a better performance on the base of that learning. This self-learning ability motivated the world to make software that operate by itself instead of wasting human energy in doing something. Breast cancer is one of the problem which can solved by self-learning of machines. Neural network is most intelligent tool design for automation. Neural is trained on some input data on the basis of this when we use this neural network on some data (related to data on which neural network is trained) then neural network tried his best to given best accuracy by using different weights and activations functions.

Introduction:

Breast cancer is caused by the development of malignant cells in the breast. The malignant cells originate in the lining of the milk glands or ducts of the breast (ductal epithelium). Breast cancer is the most common cancer in women, but it can also appear in men [1].

Signs of breast cancer:

Signs of breast cancer may include a lump in the breast, a change in breast shape, dimpling of the skin, fluid coming from the nipple, or a red scaly patch of skin [2].

Cases and deaths in 2015:

- In 2015, an estimated 231,840 new cases of invasive breast

Cancer will be diagnosed among women, as well as an estimated 60,290 additional cases of in situ breast cancer

- In 2015, approximately 40,290 women are expected to die from

Breast cancer. Only lung cancer accounts for more cancer deaths in women.

- In 2015, about 2,350 men will be diagnosed with breast cancer and 440 men will die from the disease. [3]

Background:

Classification of breast cancer:

The major categories are the histopathological type, the grade of the tumor, the stage of the tumor, and the expression of proteins and genes [4].

Histopathological: Despite the fact that breast malignancy has various histology, the significant dominant part of bosom diseases are gotten from the epithelium lining the ducts or lobules, and are named as mammary ductal carcinoma.

Grade:

Concentrates on the presence of the breast malignancy cells contrasted with the presence of ordinary breast tissue.

Stage: It depends on the measure of the growth of cancer where it initially began in the body and the areas to which it has travelled.

Neural Network Modeling:

History:

In 1943, neurophysiologist Warren McCulloch and mathematician Walter Pitts wrote a paper on how neurons might work. In order to describe how neurons in the brain might work, they

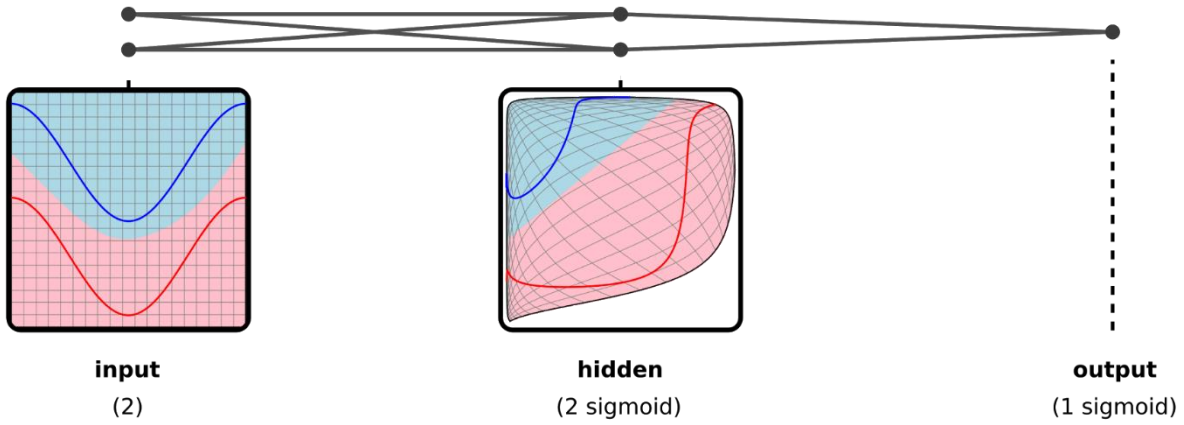
modeled a simple neural network using electrical circuits. The first multilayered network was developed in 1975, an unsupervised network [5].

Neuron:

An artificial neuron is a numerical function that considered as a model of organic neurons, a neural network. Neurons are basic units in an artificial neural system. The artificial neuron gets at least one data sources and aggregates them to create an output. Generally each input data is independently weighted, and the aggregate is gone through a non-linear function known as an activation function. Activation function most probably is a sigmoid shape.

Neural network:

A computing framework that is intended to recreate the way the human brain breaks down and process the data. Following figure is showing a neural network.



[6]

As shown from the above figure there are three major parts of a neural network one input layer, hidden layers(there can be one or many hidden layers in a neural network) and one output layer. Neural network is made up of many neurons.

Parts of a neural network:

Input layer:

E*ach unit (neuron) of the input layer, passes its assigned value to each neuron of the first hidden layer.*

Hidden layer:

The hidden layer's activity is to change the inputs into something that the output layer can utilize. At that point, each hidden layer neuron multiplies each of the incentive with its weight vector, aggregates the multiplied values, applies its activation function to this sum

(logsig, tansig and purlin) and return the esteem figured by the activation function to the next layer *(this next layer may be a next hidden layer or may be it would be output layer). If weights of a neural network change it mean that neural network has learnt something.*

Output layer:

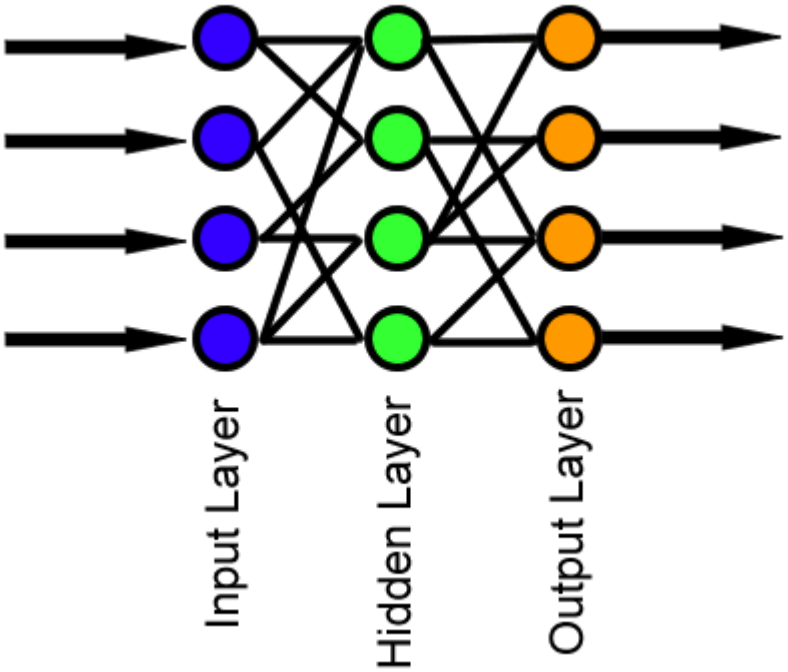
The output layer changes the concealed layer actuations into whatever scale you needed your yield to be on.

Types of a neural network:

There are many types of neural network like Feedforward, Regulatory feedback, Radial basis function, recurrent neural network, modular, physical, Dynamic, memory networks, but here I will explain feedforward and recurrent neural network.

Feedforward neural network:

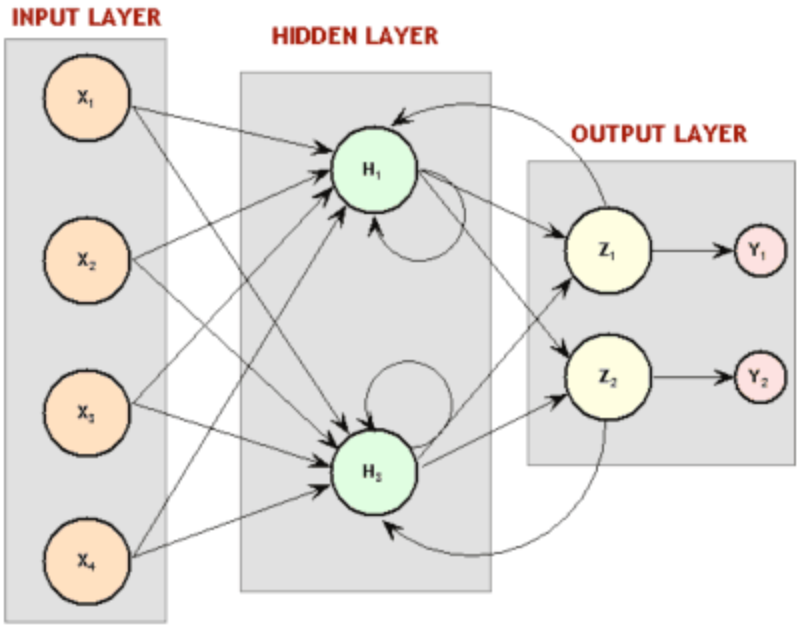
The feedforward neural network was the first and least difficult sort. It is unidirectional. In this network the data moves just from the input layer specifically through any hidden layers to the output layer without cycles/circles. Feedforward systems can be built with different sorts of units, such as binary McCulloch-Pitts neurons, perceptron, continuous neurons and sigmoidal activation.



[7]

Feedback neural network (recurrent neural network):

Loops and cycles are allowed in this neural network. Feedback neural networks, or recurrent neural networks, do contain cycles. The feedback cycles can speak to an interior state for the system that can make the network's behavior change after some time according to their inputs.



[8]

Main Part:

This part contains the methodology we develop, and justification of design decisions. Here in this section I will elaborate completely the procedure from gathering data to processing and training.

Data gathering:

We utilized a dataset accessible at UCI Machine Learning dataset repository. This informational data has 11 columns in all out first was id of patient remaining 9 columns except last one, were characteristic of patients and the last column has just two values(2 and 4), 2 demonstrates benign and 4 shows malignant. There were total 699 rows in the data. There were additionally some missing values in the data.

Initial steps:

Data downloaded from UCI machine learning dataset repository was put into a folder and in Matlab we set the path of that folder in which data is present. But before reading data from Matlab there was a problem that some of the rows in the given data contain '?'. So I just simply remove these rows from the data and then read data through Matlab using function "importdata". As patient id column (first column) was useless so I just exclude this column (1^{st}) while reading data. Data result contain values of 2 for benign and 4 for malignant, but I was using the "tansig" function which has a range between -1 to 1. Therefore, I just replace 2 with -1 and 4 with 1 in data. And then read the data again through matlab.

Training Neural Network:

A neural network was created using "newff" function and I set different attribute of learning like activation function, goal and epochs etc. After completing the stages, network architecture has been send to train function to train a neural network. Once neural network trained then to test our neural network, we use "sim" function.

Steps after training network:

To check the accuracy, I initialize a "counter" variable which gives me number of counts in which my hypothetical value match with actual value. So I used following formula to produce accuracy percentage.

Percentage of accuracy = (counter/total testing data)*100;

Hypothesis, Result and analysis:

First hypothesis:

Statement:

Our result will be improved and will be better than previous result by increasing the number of nodes in hidden layer and using "tansig" as an activation function. During experiment goal and epochs were equal to 0.01 and 100 respectively.

Experiments:

EXP No.	Hidden Layer	Result (%)
1	10	98.0296
2	20	98.5222
3	30	98.5222
4	40	99.0148
5	50	98.5222

Result:

As shown from the experiments that by increasing the nodes in the hidden layer result is not constantly increasing so our hypothesis is wrong but 4th experiment shows that at 40 (size of hidden layer) result is maximum.

Second hypothesis:

Statement:

If we take bigger data in input during training and train network with "trainb (Batch training with weight and bias learning rules)" and use "tansig" as activation function then our neural network performance and accuracy will go down.

Experiments:

EXP No.	Input Data	Result (%)
1	60	90.0482
2	120	94.4938
3	180	96.4215
4	240	95.5414

5	300	96.2848
7	360	95.8175
8	420	97.5369

Result:

As shown from the above experiments accuracy sometime decreases and sometime increases, therefore, our hypothesis goes wrong but still we got maximum result at 420 which is maximum input data as compare to all other inputs.

Third Hypothesis:

Statement:

If we train our neural network on "trainlm" function instead of "train" then we will never get a better accuracy result for small input data as compare to big data because "trainlm" is used for big data. So if we apply "trainlm" on complete data then we will get best accuracy result.

Experiments:

EXP No.	Input data	Results
1	120	92.8952
2	240	96.3883
3	360	96.9040
4	480	98.0296
5	600	99.3007

Results:

Experiments shows that our accuracy is going to increase as we are increasing the input data.

Fourth Hypothesis:

Statement:

By increasing the number of hidden layer and increasing the nodes in each hidden layer then the accuracy and performance will be increased.

Experiments:

We are performing all the experiments using 480 as an input data.

EXP No.	No. of hidden layers	No. of nodes	Result
1	1	20	98.5222
2	2	20	98.0296
		40	

3	3	20	99.5074
		40	
		60	
4	4	20	97.5369
		40	
		60	
		80	

Results:

As clear from the table that our result is increasing, therefore, our hypothesis is not correct but we got high performance at 3rd experiment, it mean these are the setting where we can get better performance.

Conclusion:

Neural Networks are now being utilized as a part of the characterization of breast malignancy. The purpose of this exercise to explore some methods through which enhance the execution and exactness of neural network to characterize breast malignancy. We can improve performance by many parameters but here we discuss some like hidden layers, weight, activation functions. As seen in experiment that which setting can give us better performance during training of a neural network, therefore, these tests will help a lot while investigation breast cancer.

Refrences:

[1]: https://www.nature.com/subjects/breast-cancer

[2] https://en.wikipedia.org/wiki/Breast_cancer

[3]: https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/breast-cancer-facts-and-figures/breast-cancer-facts-and-figures-2015-2016.pdf

[4]: https://en.wikipedia.org/wiki/Breast_cancer_classification

[5]: https://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/History/history1.html

[6]: http://colah.github.io/posts/2015-09-NN-Types-FP/

[7]: https://en.wikipedia.org/wiki/Feedforward_neural_network

[8]: https://www.google.com/search?q=feedback+neural+network+pictures&client=firefox-b-ab&tbm=isch&source=iu&ictx=1&fir=34oQDPHpFtwpsM%253A%252C3z2iBk9WK3UYGM%252C_&usg=__Gm7J6NIYSfYbTr3gqHpdA8IZQp0%3D&sa=X&ved=0ahUKEwit88mtgeLXAhVMN48KHfqrB84Q9QEIMzAE&biw=1366&bih=656#imgrc=_