EE 550

Artificial Neural Networks

Homework 3

Sezen Perçin – 2018401000

For all the questions below, I included the weights used for the biasing purposes to the weight matrix and uploaded them with their own updating rule implemented in the script accordingly.

1.For the first question I have set the inputs and the corresponding desired output values first and initialized the weight matrices. The learning rate is hosen as 0.6 and the MLP structure chosen can be seen below (Fig.1.1). Then until the defined total error variable “errp” is less than equal to 0.001, for each sample in the sample set, a forward operation is performed to find the error value at the output.

This value is used for the calculation of the del3 at the output and del values of the hidden layers through the usage of del3 and the backpropagation algorithm. For each sample next, the amount of change that will be made on the weight matrices is calculated and added to the weight matrices. Total error vs number of epoch plot (Fig.1.2) and the actual outputs of the inputs calculated with the resultant weight matrices (Table 1.1) can be found below.

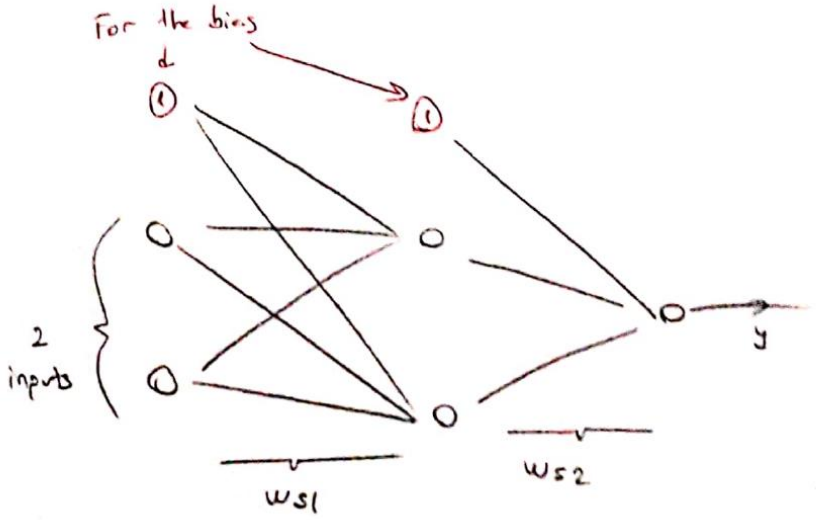

Fig.1.1

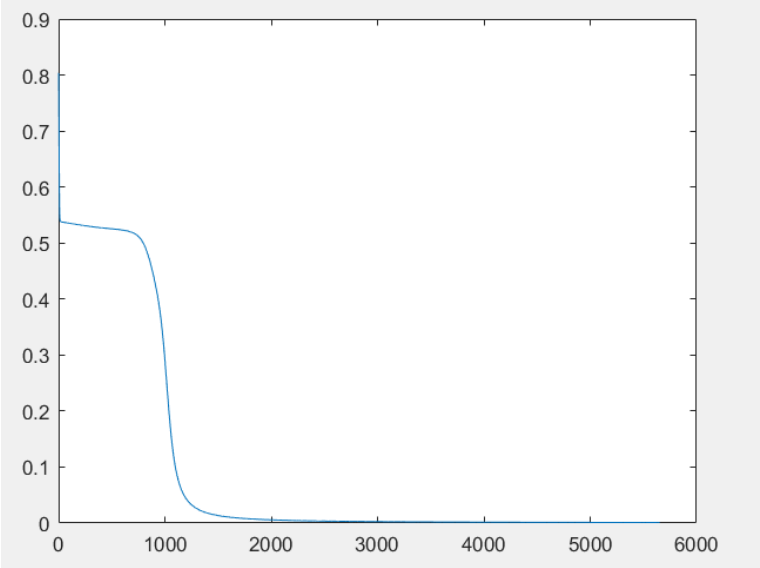

Fig 1.2

It took 5653 epochs for the total error to be less than or equal to 0.001 for the shown case.

T =

4×4 table

input1	input2	desired output	predicted output
0	0	0	0.0246761978692886
0	1	1	0.978872530415268
1	0	1	0.978602995796244
1	1	0	0.022059526382682

Table 1.1

2.For the second question first I defined a function to calculate the output values for the given function. Since sigmoid function is significant for the values between (0,1) I have normalized my input and corresponding output set. Then I initialized the weight matrices and set learning rate as 0.7. The MLP I have used can be seen below (Fig 2.1). Then until the define total error value ”errp” is less than or equal to the chosen threshold 0.1 ,for each sample I performed a forward operation to find the error at the output.

Then this error value is used for the calculation of del4 values and the other del values through the usage of backpropagation algorithm and del4.The amount of changes (delta values) that should be made on the weight matrices is found and added to the weight matrices. It took 470

epoch for errp value to be less than or equal to 0.1. As a result, the actual outputs(shown as red circles) and desired values (shown as blue x marks) plotted Fig2.3 (only 25 samples) and total error vs. number of epochs Fig.2.2 can be found below.

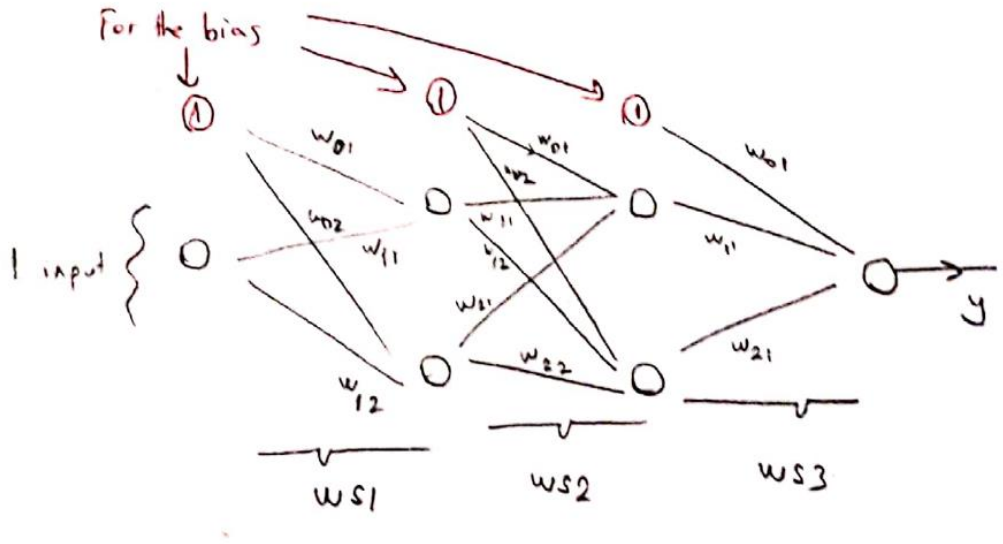

Fig. 2.1

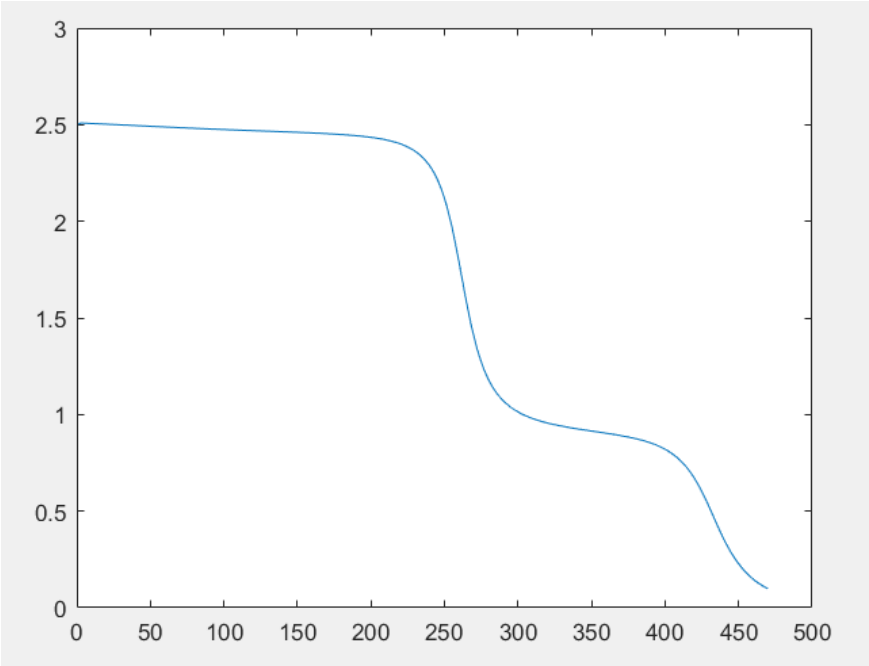

Fig. 2.2

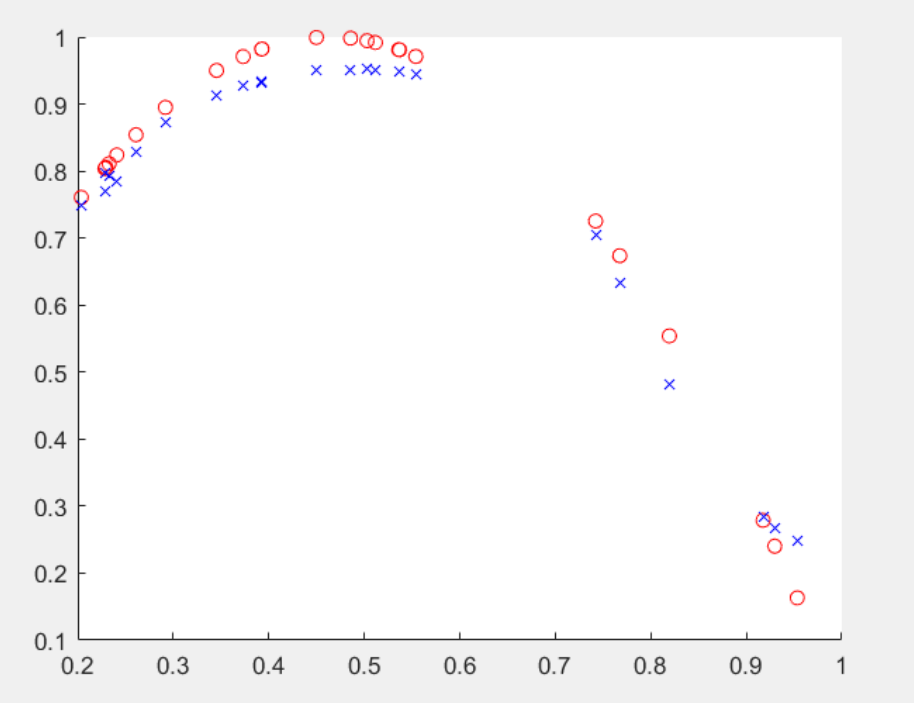

Fig.2.3

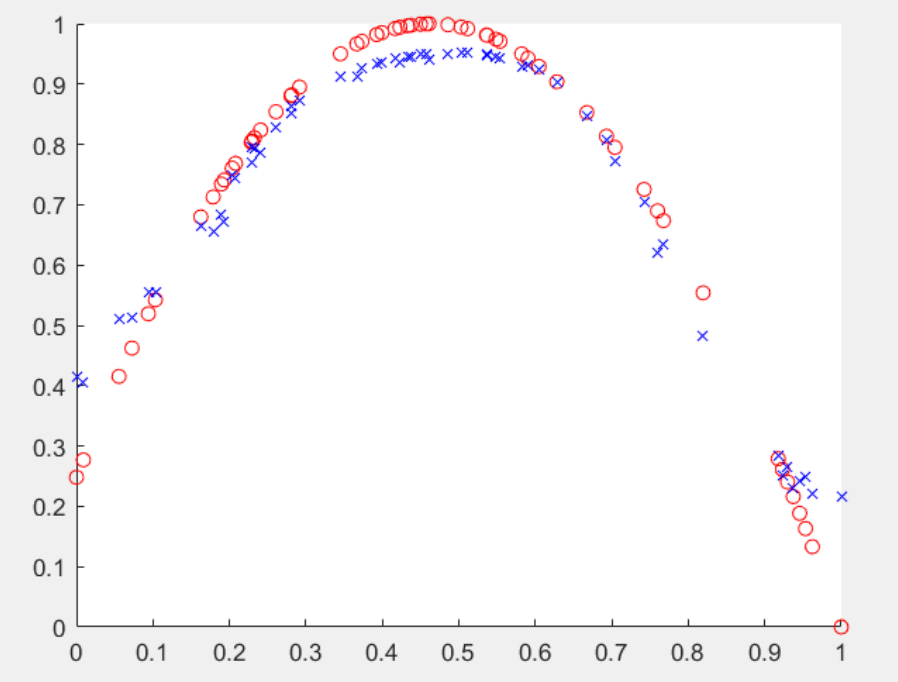

Fig. 2.4

Fig.2.4 shows all the data set plottet for the desired outputs and actual outputs.

3.For the third question, I have obtained the data set from the https://archive.ics.uci.edu. Since they will go into the sigmoid function I normalized them by mapping the input values to range (0,1). For this operation I have used the MLP that can be seen in the Fig.3.1. Next step was the classification of the input samples so that I can take 100 samples from each class and form the sample set. After the initialization of the weight matrices I set learning rate to 0.7 and specified that the iterations should stop when the total error is less than or equal to 0.1 (total error variable is defined as “errp” in the script).

For each sample in the sample set first I have performed a forward operation to get the error that is defined as (desired output-actual output). Then I have used this value to find del3 value and by using backpropagation I found other values. Those del values later used for the calculation of the amount of change that must be made in the weight matrices. It took 89275 epochs for the total error to be equal or less than 0.1.

As the final step the total error vs number of epochs is plotted Fig. 3.2 and the desired output values (showed as red circles on the plot) and actual output values (showed as blue x marks on the plot) are shown for the test values that was took from the input set (Fig 3.3).

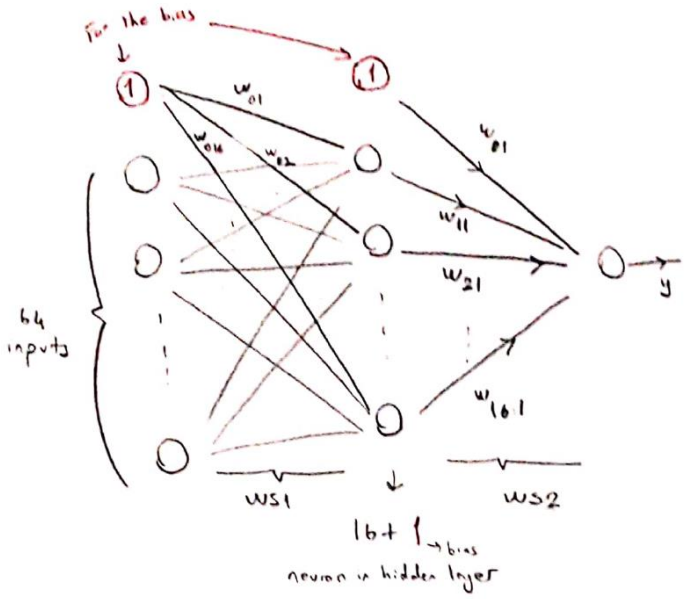

Fig 3.1

It is a 3-layer model which has 16 neuron in the hidden layers and take 64 inputs. Also in each layer except the output layer biasing is used.

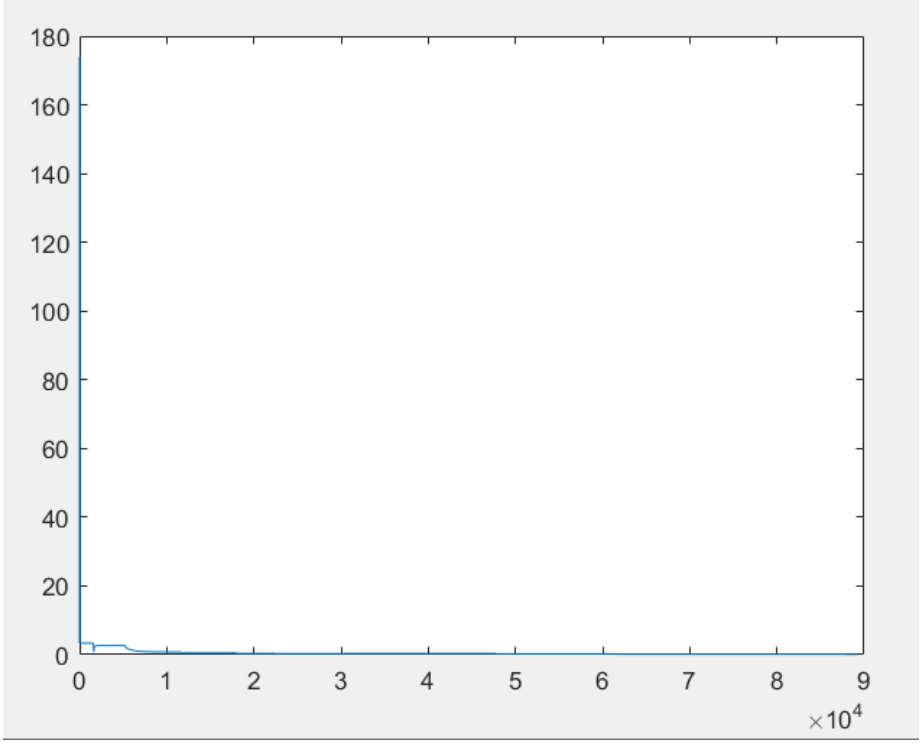

Fig.3.2

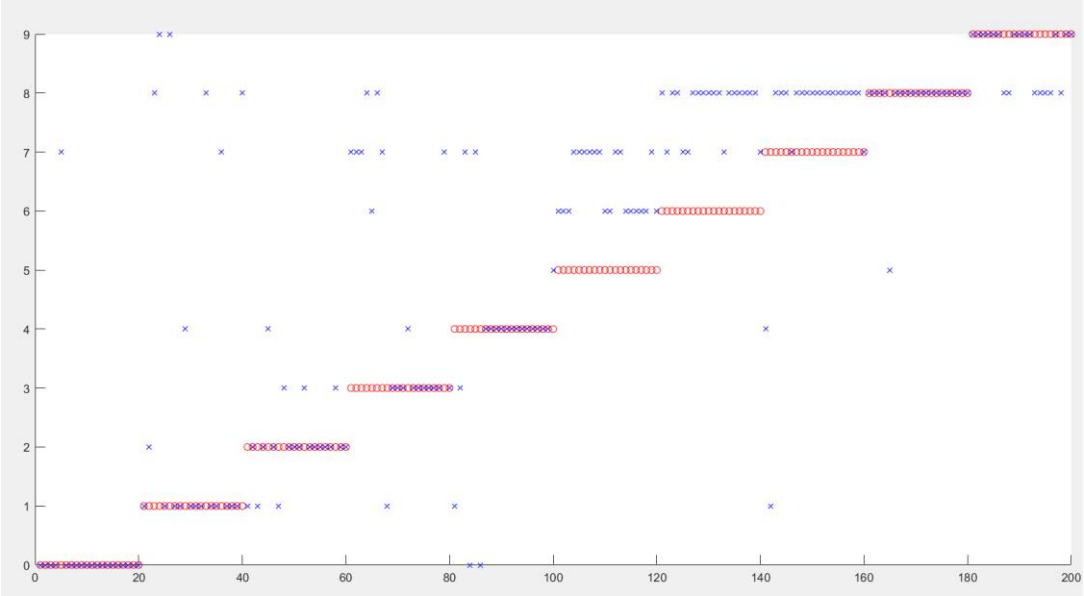

Fig.3.3

For the some of the classes the MLP works is more successful. Mostly it has a problem with the upper bound which I suspect to result from the calculation issues happening with limitations the sigmoid function.