Report

Yekta Demirci

25/03/2018

1 PART 1

1.1 Network 1 Graph

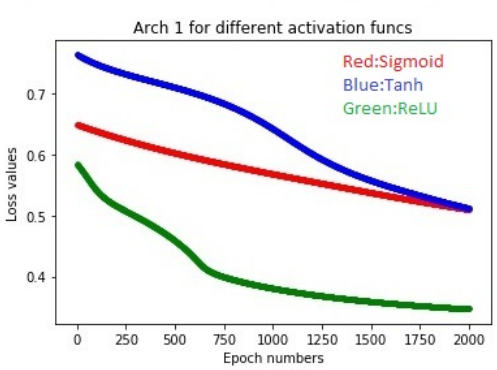

Figure 1: Loss value versus epoch for arch1

1.2 Network 2 Graph

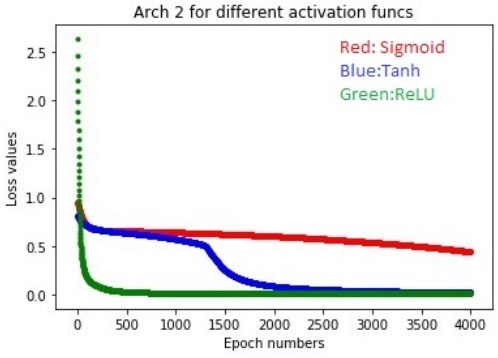

Figure 2: Loss value versus epoch for arch2

1.3 Network 3 Graph

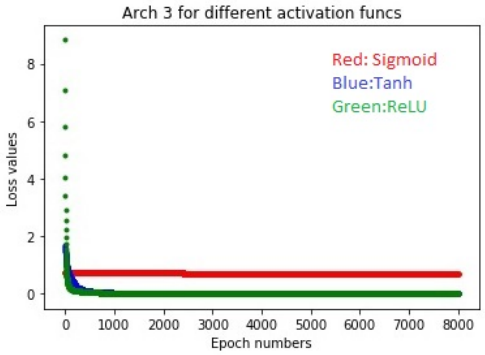

Figure 3: Loss value versus epoch for arch3

1.4 Final Accuracy For Each Network

Network1	0.995633
Network2	1.0
Network3	1.0

1.5 Hyperparameter Optimization

1.5.1 Optimization

Nested if else statements are used 3 times for each sigmoid,tanh and relu combination for each learning rate that results with 270 combinations. Test value for the best case is 1

S is `sigmoid`, T is `tanh` and R is `relu` activation function.

Layer Activations	Learning Rate									
	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
SSS	0.624454	0.995633	1.0	0.995633	1.0	0.995633	1.0	1.0	1.0	1.0
SST	0.975983	1.0	1.0	1.0	1.0	0.995633	1.0	1.0	1.0	1.0
SSR	0.908297	0.995633	1.0	1.0	1.0	1.0	1.0	0.9869	1.0	1.0
STS	0.984716	1.0	1.0	0.934498	1.0	1.0	1.0	1.0	1.0	1.0
STT	0.709607	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
STR	0.995633	0.995633	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
SRS	0.744541	0.9869	0.995633	1.0	1.0	1.0	1.0	1.0	1.0	1.0
SRT	0.995633	0.694323	0.997817	1.0	1.0	1.0	1.0	0.554585	1.0	1.0
SRR	0.554585	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.99345	1.0
TSS	0.980349	0.995633	1.0	1.0	1.0	0.995633	1.0	1.0	1.0	1.0
TST	0.847162	1.0	1.0	1.0	0.997817	1.0	1.0	1.0	1.0	1.0
TSR	0.997817	1.0	0.997817	1.0	1.0	1.0	1.0	1.0	1.0	1.0
TTS	0.997817	0.997817	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
TTT	1.0	1.0	1.0	1.0	0.99345	1.0	1.0	1.0	1.0	1.0
TTR	1.0	1.0	0.99345	0.99345	1.0	1.0	0.997817	1.0	1.0	1.0
TRS	0.978166	0.997817	0.995633	1.0	1.0	0.997817	1.0	1.0	1.0	1.0
TRT	0.995633	1.0	0.1	1.0	1.0	1.0	1.0	0.997817	1.0	1.0
TRR	0.995633	0.997817	0.997817	0.997817	1.0	1.0	1.0	1.0	0.997817	1.0
RSS	0.967249	0.99345	0.1	0.997817	0.995633	1.0	1.0	1.0	1.0	1.0
RST	0.901747	1.0	1.0	1.0	1.0	1.0	0.99345	1.0	1.0	1.0
RSR	0.995633	1.0	1.0	0.99345	1.0	1.0	1.0	1.0	1.0	1.0
RTS	0.980349	1.0	1.0	0.997817	1.0	0.995633	1.0	1.0	1.0	0.99345
RTT	0.984716	0.980349	1.0	0.99345	1.0	1.0	1.0	1.0	0.997817	0.997817
RTR	1.0	1.0	1.0	0.997817	1.0	1.0	0.997817	1.0	1.0	0.997817
RRS	0.997817	0.995633	0.995633	1.0	1.0	0.99345	1.0	1.0	1.0	1.0
RRT	0.949782	0.997817	1.0	1.0	0.997817	1.0	1.0	0.997817	0.99345	1.0
RRR	0.991266	1.0	1.0	1.0	1.0	0.997817	1.0	1.0	1.0	1.0

1.5.2 Training and Test

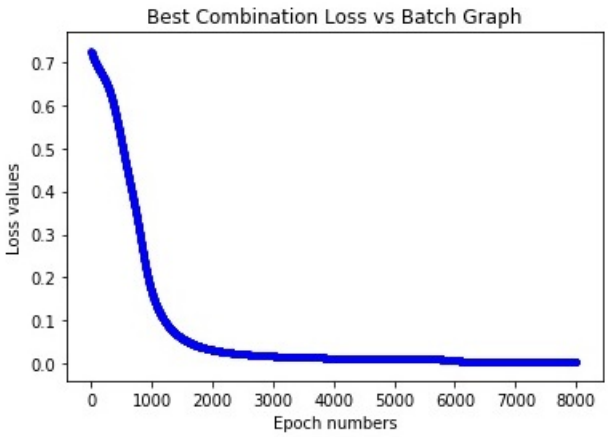

Figure 4: Training loss values versus epoch(*Test accuracy is 1.00 for the best training accuracy constants)

2 PART 2

In this part, 5 hidden layers are implemented. Because from previous parts, it is observed that as the hidden layers are increased,better results are found. Then 20,15,10,6,4 nodes are implemented for respected hidden layers. As ac-

tivation function tanh is used for the first 4 hidden layers and sigmoid for the last one,because it can be seen from optimization table for all learning rates TT and another activation function gave 1.00 accuracy value which tells that, tanh functions seems the best with combining with another activation function at some layer, eventhough in the first part, reLU gave better accuracy. Epoch is used as 5000 because after some point, loss value decrease is so low,therefore no need to keep training it.

As a result 0.728 accuracy is found from test inputs.

Respected training loss vs epoch graph is given below

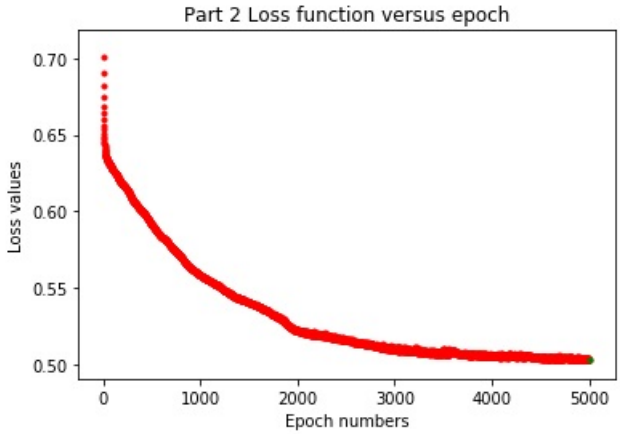

Figure 5: Loss value versus epoch for part2