

# Exercise 1 – Neural Networks

Q1.

- i.  $a = \sum_{i=1}^n w_i x_i$   
 $a = w_1 x_1 + w_2 x_2 + w_3 x_3$   
 $a = 0.2 \times 1 + (-1) \times 0 + 0.5 \times 1$   
 $a = 0.2 + 0 + 0.5$   
 The activation is  $a = 0.7$
- ii.  $y = \begin{cases} 1, & a \geq h \\ 0, & a < h \end{cases}$   
 $a = 0.7 > h = 0.5 \Rightarrow$  The output is  $y = 1$
- iii.  $a = 0.7 < h = 0.8 \Rightarrow$  The output is  $y = 0$

Q2.

- i.  $a = -0.2 < h = 0.5 \Rightarrow y = 0$
- $w'_1 = w_1 + \alpha(t - y)x_1 = 0.1$   
 $w'_2 = w_2 + \alpha(t - y)x_2$   
 $w'_2 = -1 + 0.1(1 - 0) \times 0$   
 $w'_2 = -1$   
 $w'_3 = w_3 + \alpha(t - y)x_3$   
 $w'_3 = 0.5 + 0.1(1 - 0) \times 0$   
 $w'_3 = 0.5$
- ii.  $\vec{w}' = \langle 0.1, -1, 0.5 \rangle$
- iii.
- Pattern  $p_2$ :  
 $a = \sum_{i=1}^n w_i x_i = 0.1 \times (-1) + (-1) \times 0 + 0.5 \times 0 = -0.1$   
 $a = -0.1 < h = 0.5 \Rightarrow y = 0$
- $w''_1 = w'_1 + \alpha(t - y)x_1$   
 $w''_1 = 0.1 + 0.1(1 - 0) \times (-1)$   
 $w''_1 = 0$   
 $w''_2 = w'_2 + \alpha(t - y)x_2$   
 $w''_2 = -1 + 0.1(1 - 0) \times 0$   
 $w''_2 = -1$   
 $w''_3 = w'_3 + \alpha(t - y)x_3$   
 $w''_3 = 0.5 + 0.1(1 - 0) \times 1$   
 $w''_3 = 0.6$
- $\Rightarrow \vec{w}'' = \langle 0.1, -1, 0.5 \rangle$

- Pattern p<sub>3</sub>:

$$a = \sum_{i=1}^n w_i x_i = 0.1 \times (-1) + (-1) \times 0 + 0.5 \times 1 = 0.4$$

$$a = 0.4 < h = 0.5 \Rightarrow y = 0$$

$$w_1''' = w_1'' + \alpha(t - y)x_1$$

$$w_1''' = 0 + 0.1(1 - 0) \times (-1)$$

$$w_1''' = -0.1$$

$$w_2''' = w_2'' + \alpha(t - y)x_2$$

$$w_2''' = -1 + 0.1(1 - 0) \times 1$$

$$w_2''' = -0.9$$

$$w_3''' = w_3'' + \alpha(t - y)x_3$$

$$w_3''' = 0.6 + 0.1(1 - 0) \times 0$$

$$w_3''' = 0.6$$

$$\Rightarrow \vec{w}''' = \langle -0.1, -0.9, 0.6 \rangle$$

- Pattern p<sub>4</sub>:

$$a = \sum_{i=1}^n w_i x_i = -0.1 \times (-1) + (-0.9) \times 1 + 0.6 \times 1$$

$$a = 0.1 - 0.9 + 0.6$$

$$a = -0.2$$

$$a = 0.4 < h = 0.5 \Rightarrow y = 0$$

$$w_1'''' = w_1''' + \alpha(t - y)x_1$$

$$w_1'''' = -0.1 + 0.1(0 - 0) \times (-1)$$

$$w_1'''' = -0.1 + 0$$

$$w_1'''' = -0.1$$

$$w_2'''' = w_2''' + \alpha(t - y)x_2$$

$$w_2'''' = -0.9 + 0.1(0 - 0) \times 1$$

$$w_2'''' = -0.9 + 0$$

$$w_2'''' = -0.9$$

$$w_3'''' = w_3''' + \alpha(t - y)x_3$$

$$w_3'''' = 0.6 + 0.1(0 - 0) \times 1$$

$$w_3'''' = 0.6 + 0$$

$$w_3'''' = 0.6$$

$$\Rightarrow \vec{w}'' = \langle -0.1, -0.9, 0.6 \rangle$$

- iv. The learning does not change the weight vector for pattern p<sub>4</sub> because the target output of the pattern is  $t = 0$  and the output  $y = 0$ , therefore the update equation would become:

$$w_i' = w_i + \alpha(t - y)x_i \Rightarrow w_i' = w_i + \alpha(0 - 0)x_i \Rightarrow w_i' = w_i + \alpha \times 0 \times x_i \Rightarrow w_i' = w_i + 0$$

$$\Rightarrow w_i' = w_i$$

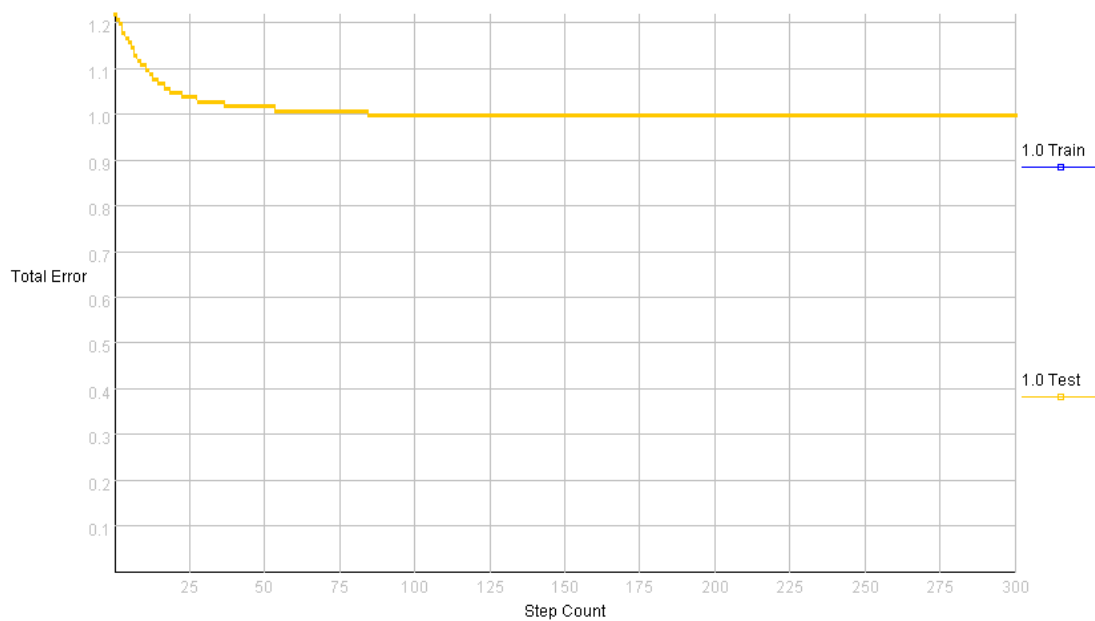
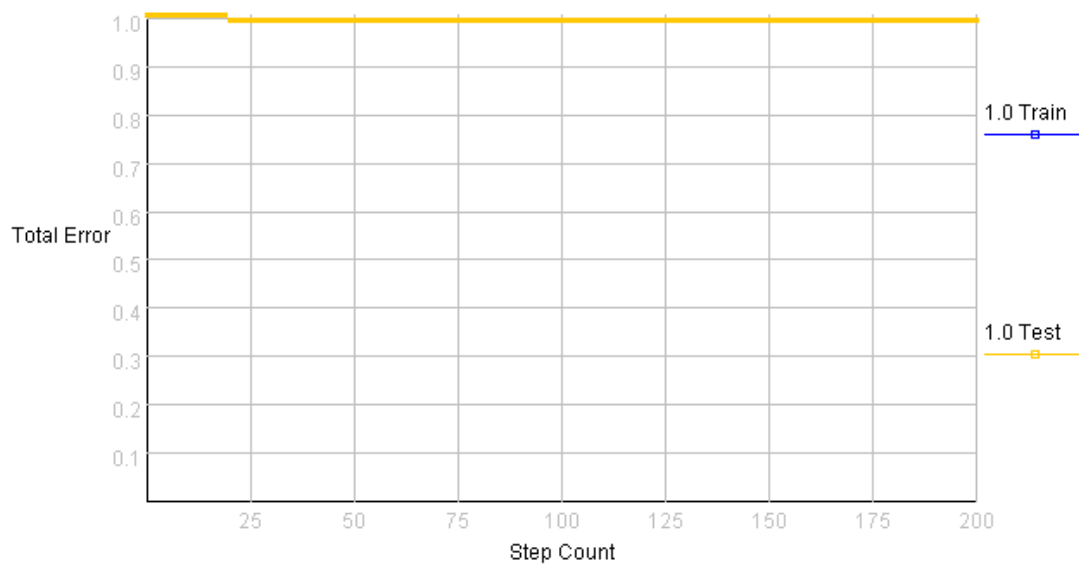
As a result, the weight vector for pattern p<sub>4</sub> is the same as the weight vector for pattern p<sub>3</sub>.

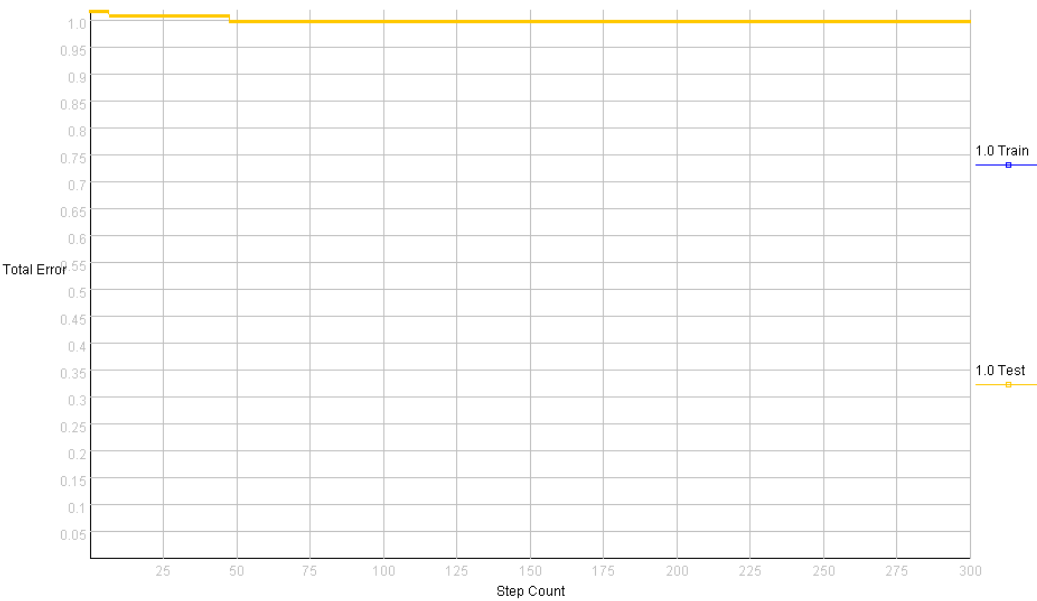
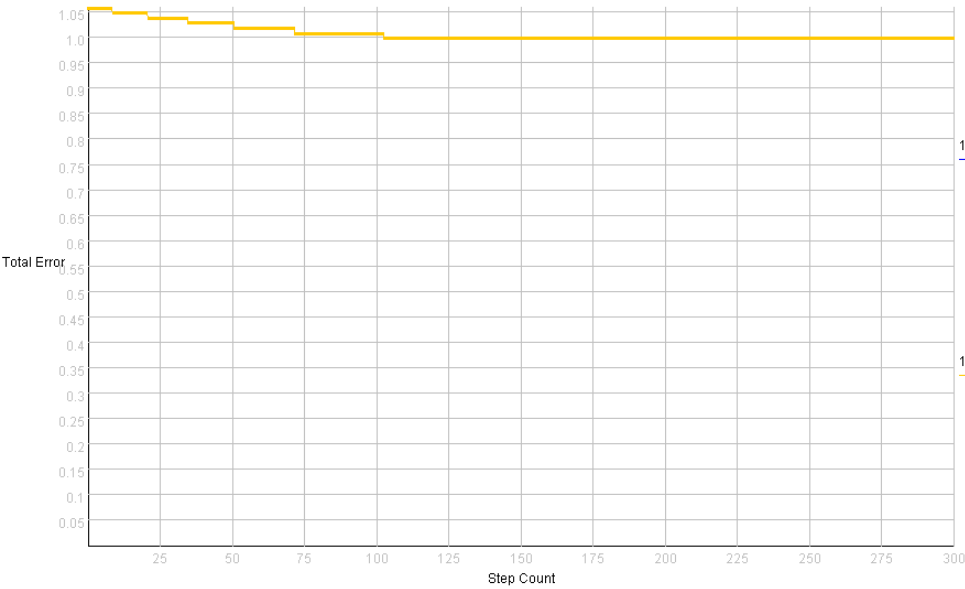
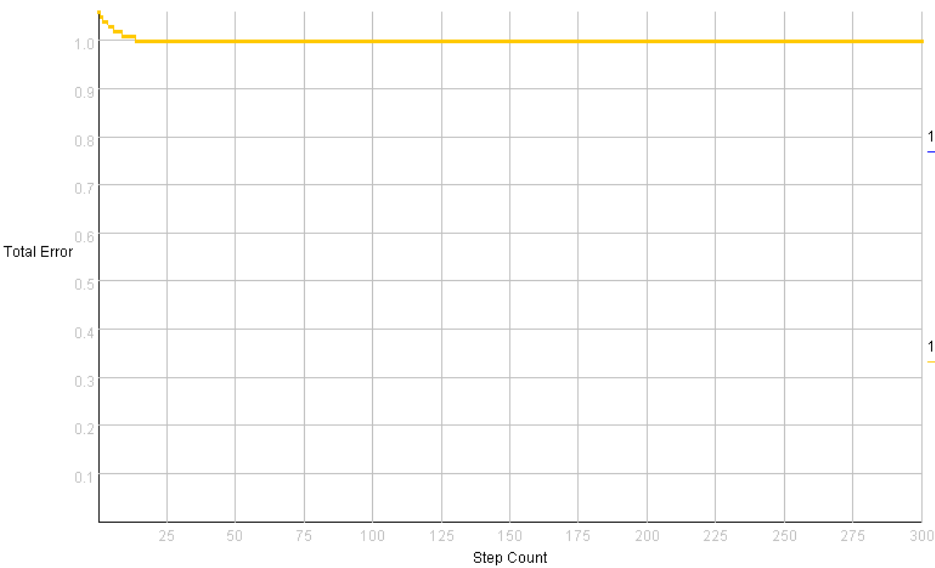
v. Yes.

Q3.

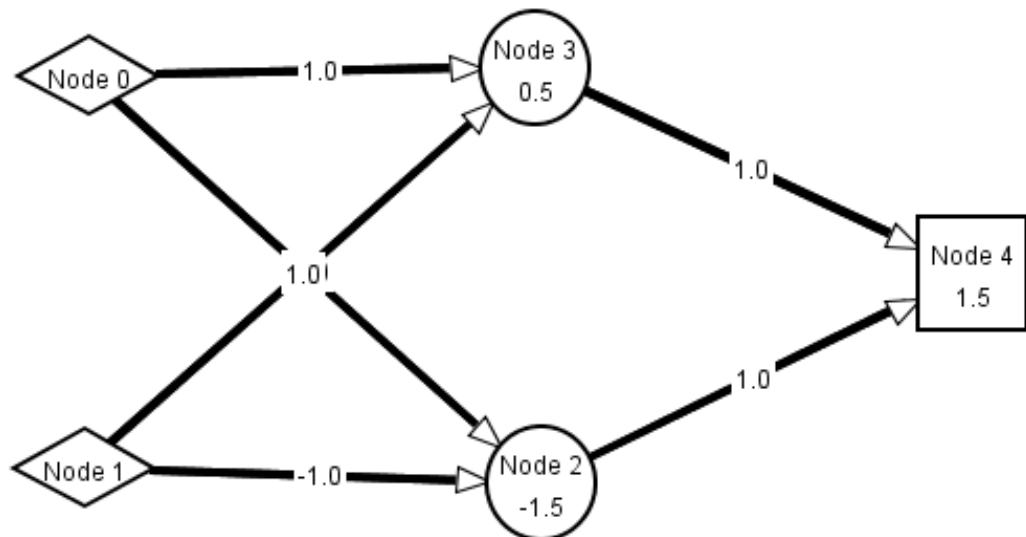
i. Initial Test Error: 1.0211

Test no	Initial $w_1$	Initial $w_2$	Final $w_1$	Final $w_2$	Correct test ex (50)	Correct test ex (100)	Correct test ex (150)	Correct test ex (200)	Correct test ex (250)	Correct test ex (300)
1	0.5	0.21	0.09	0.06	50%	50%	50%	50%	-	-
2	0.95	-0.61	0.01	-0.03	50%	50%	50%	50%	50%	50%
3	0.29	-0.01	0.0	-0.01	50%	50%	50%	50%	50%	50%
4	-0.86	0.71	-0.02	0.03	75%	75%	50%	50%	50%	50%
5	0.38	0.37	0.04	0.04	25%	25%	25%	25%	25%	25%



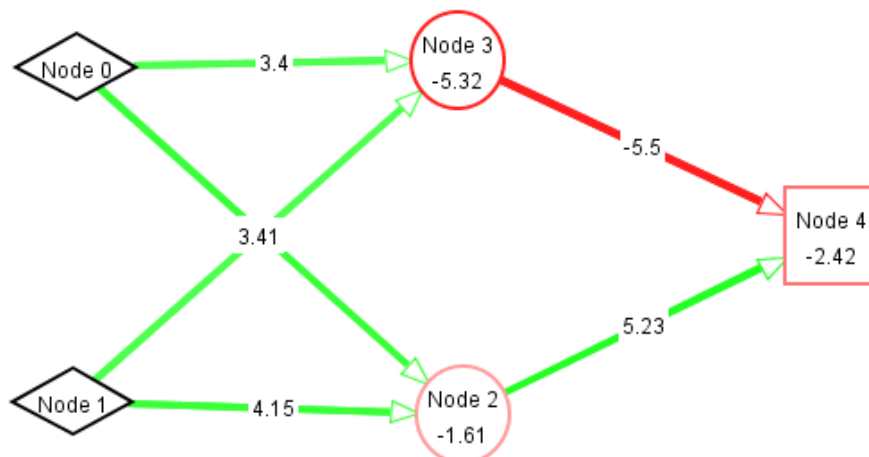


- ii. The single neuron doesn't converge to classify all test patterns correctly
- iii. XOR is not a linearly separable problem.
- iv. The network won't solve XOR.
- vi. The network that solves XOR:

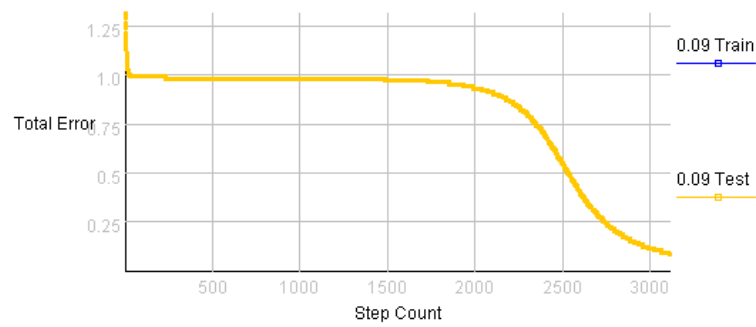


Trial 1 (weights and steps):

Reached target training error of 0.0999 after 3107 steps.  
Test Error: 0.0999

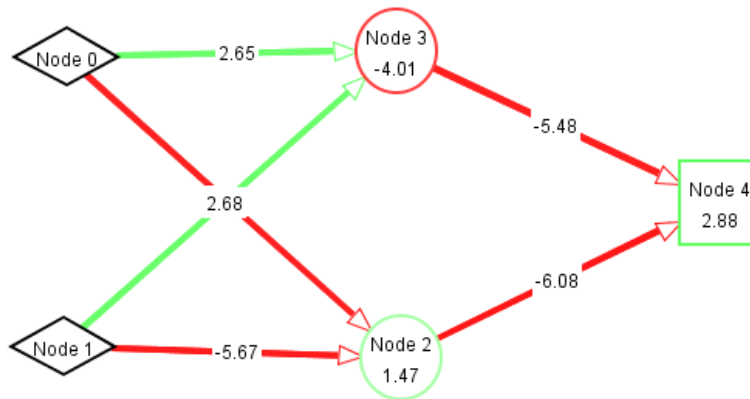


Trial 1 (error graph):



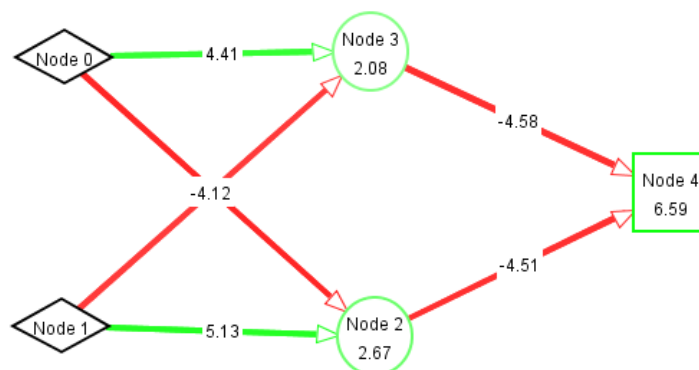
Trial 2 (weights and steps):

Reached target training error of 0.0998 after 2893 steps.  
Test Error: 0.0998

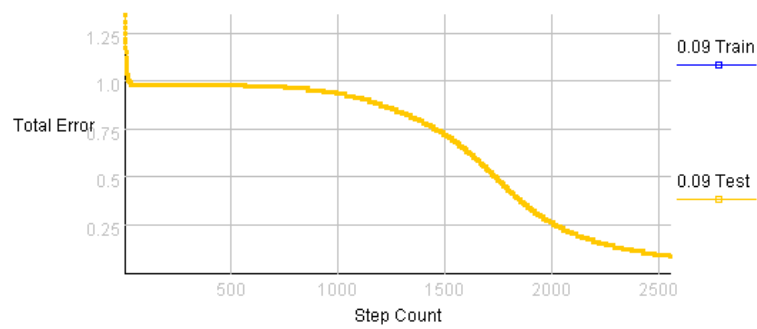


Trial 3 (weights and steps):

Reached target training error of 0.1 after 2549 steps.  
Test Error: 0.1

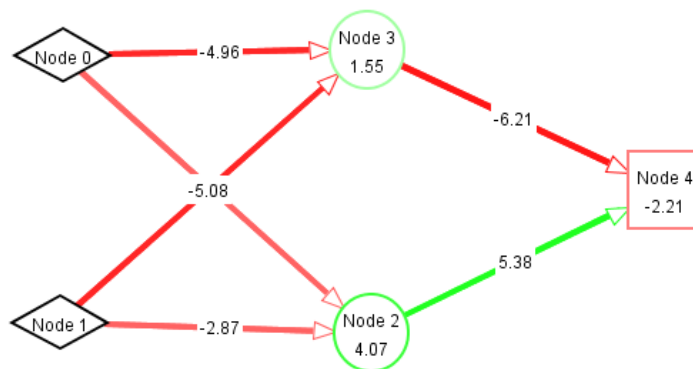


Trial 3 (error graph):



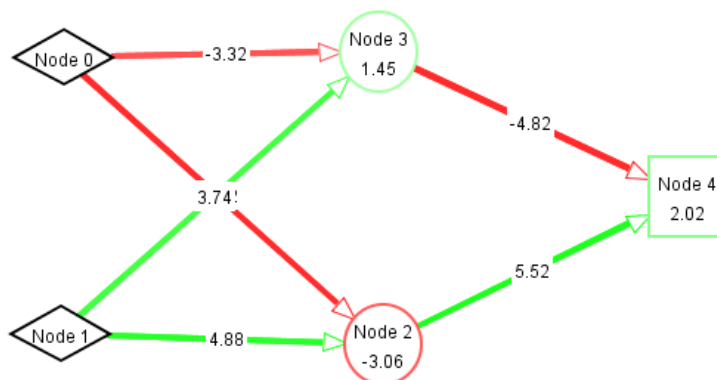
Trial 4 (weights and steps):

Reached target training error of 0.0999 after 2166 steps.  
Test Error: 0.0999

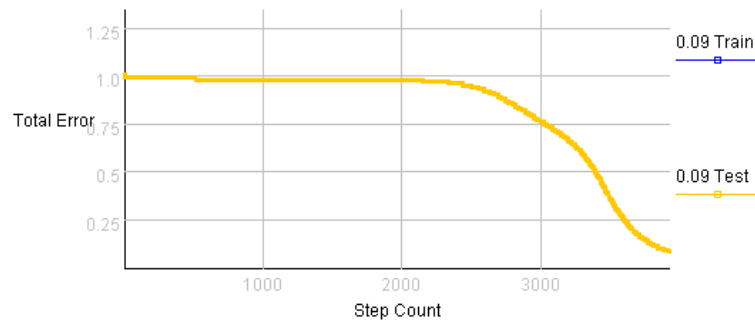


Trial 5 (weights and steps):

Reached target training error of 0.0999 after 3924 steps.  
Test Error: 0.0999

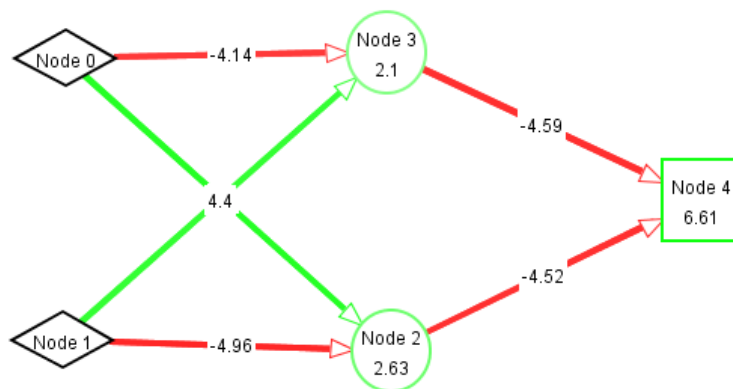


Trial 5 (error graph):



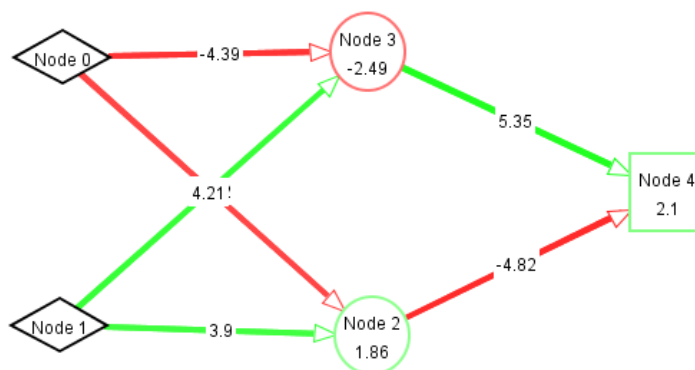
Trial 6 (weights and steps):

Reached target training error of 0.0999 after 3223 steps.  
Test Error: 0.0999



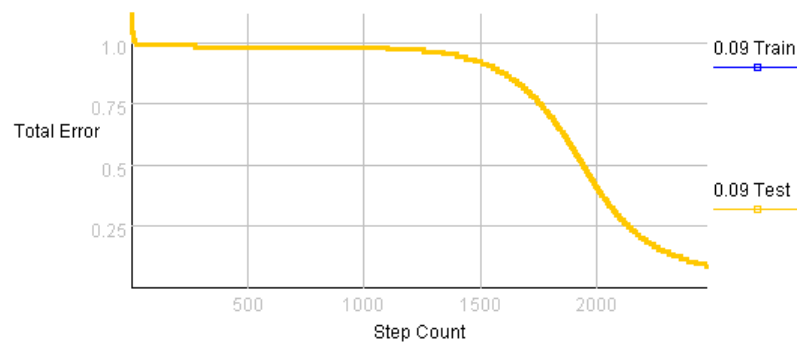
Trial 7 (weights and steps):

Reached target training error of 0.0998 after 2469 steps.  
Test Error: 0.0998



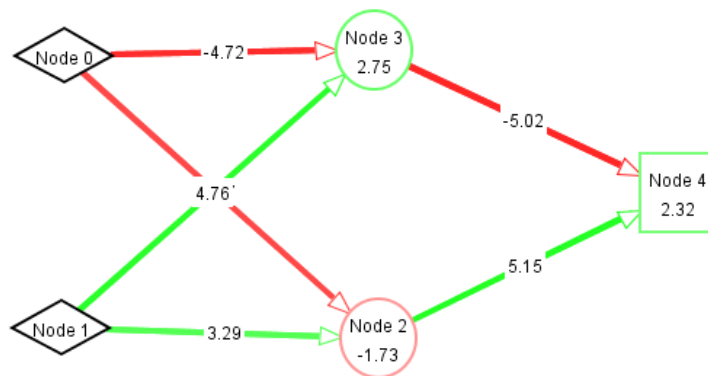


Trial 7 (error graph):



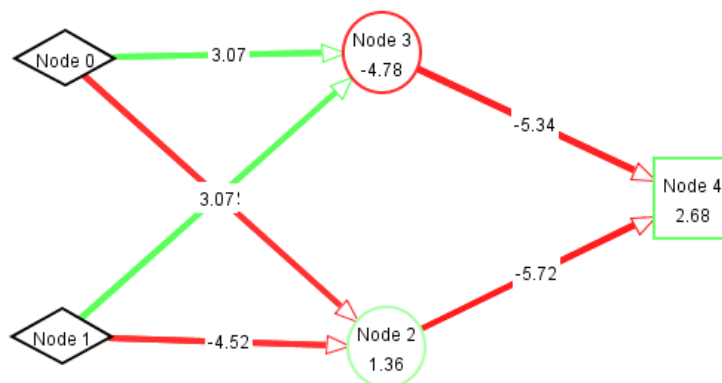
Trial 8 (weights and steps):

Reached target training error of 0.0999 after 2096 steps.  
Test Error: 0.0999

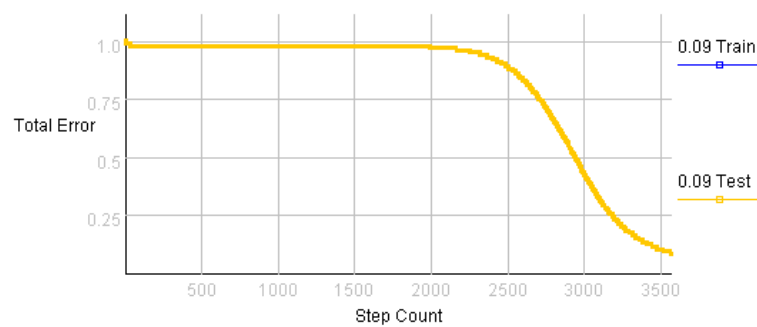


Trial 9 (weights and steps):

Reached target training error of 0.0999 after 3561 steps.  
Test Error: 0.0999

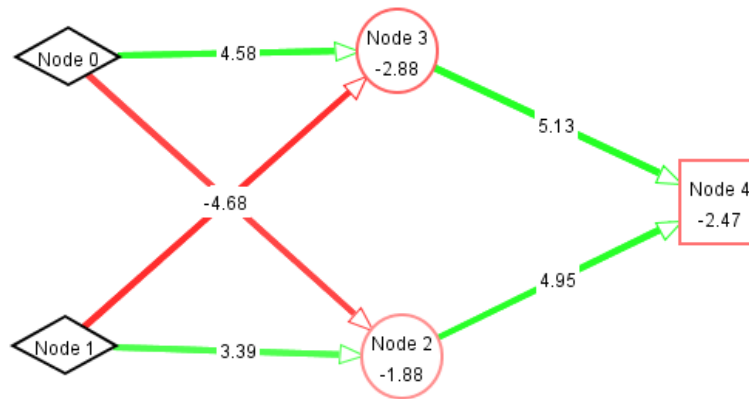


Trial 9 (error graph):



Trial 10 (weights and steps):

Reached target training error of 0.0999 after 3419 steps.  
Test Error: 0.0999



- vii. What do you notice about the error on different trials. Explain your observations.

Each trial stops at a different number of steps. Some trials fail to stop before 50,000 steps.