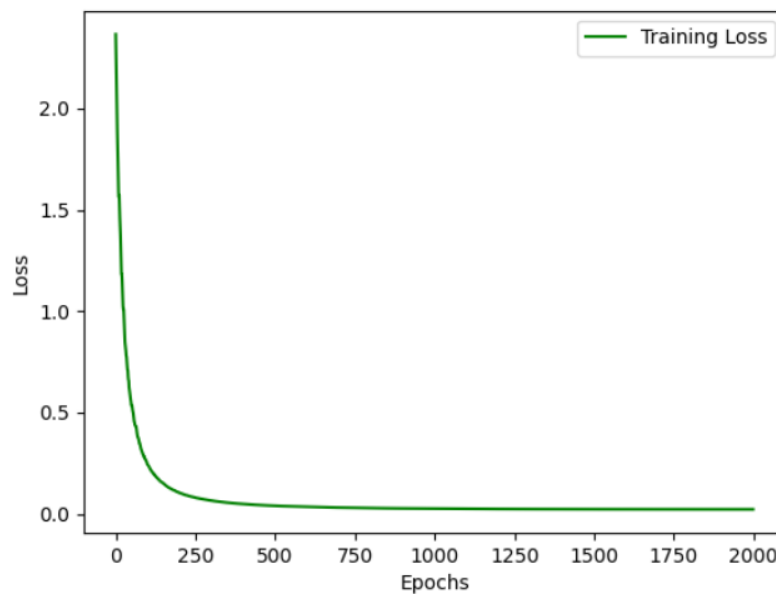


## Assignment1:

**Question1:** The max relative error  $7.047383216379126e-16$

	$W$	$b$
<b>Layer 1</b>	$3.01e-16$	$2.51e-16$
<b>Layer 2</b>	$7.05e-16$	$2.21e-16$

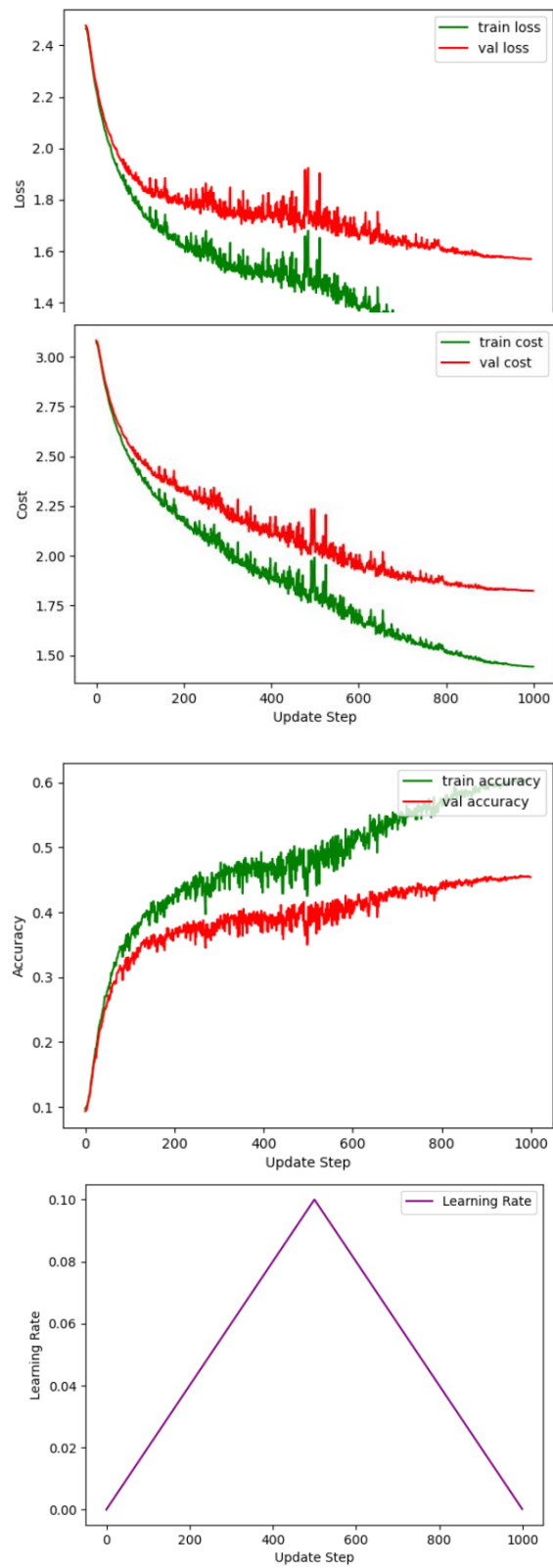


Accuracy on train data: from 0.15 – 1

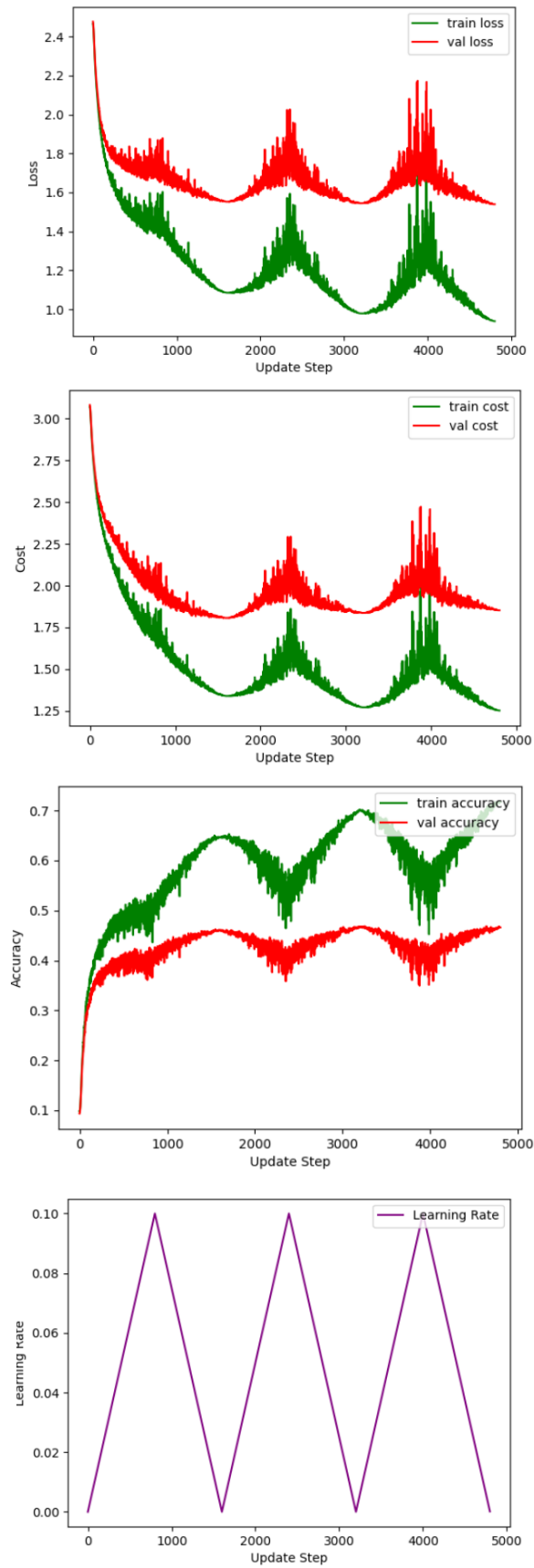
The function `relative_error` calculate the two gradients relative error. I calculate two layers gradients and return the value.

The max `relative_error` is  $7.047383216379126e-16$  less than the  $1e-6$ ,

## Question2:



$\eta_{\min} = 1e-5, \eta_{\max} = 1e-1, \text{epoch} = 10, n_s = 500.$



The  $\eta_{\min} = 1e-5$ ,  $\eta_{\max} = 1e-1$ ,  $n_s = 800$ ,  $\text{cycle} = 3$ ,  $\text{epoch} = 48$ .

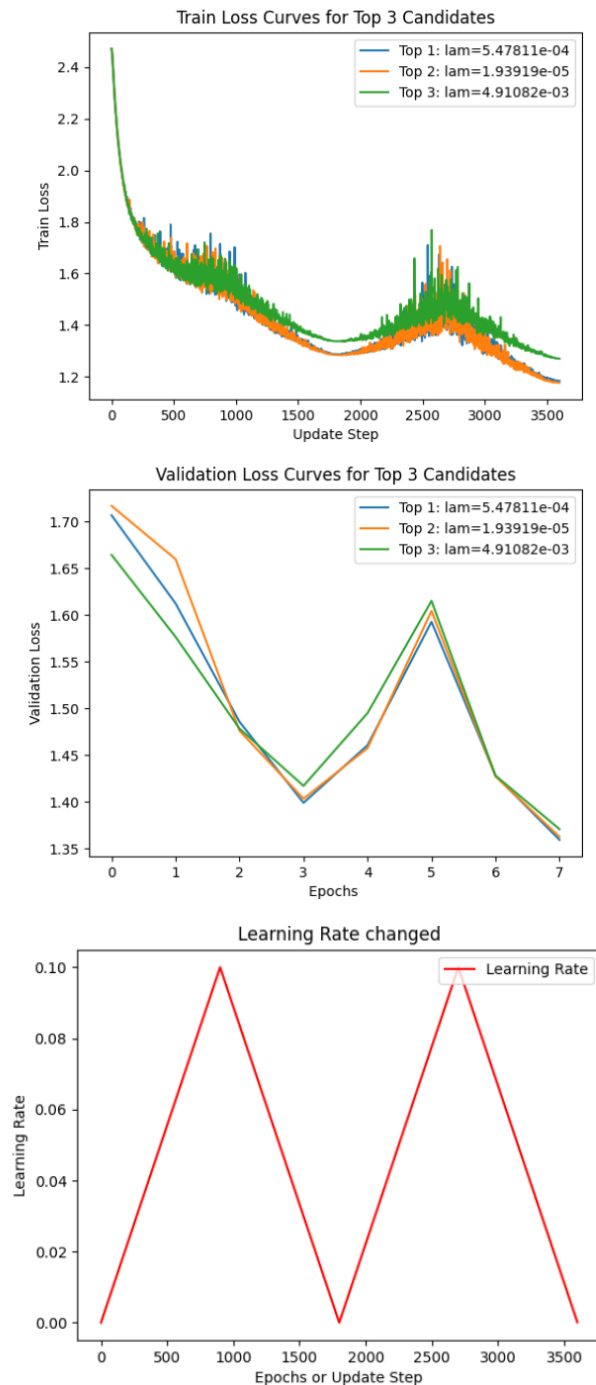
### Question3:

The range of the value I searched during the coarse search for  $\lambda$  is from  $\text{lam\_min} = 10^{-5} = 1.00\text{e-}05$  to  $\text{lam\_max} = 10^{-1} = 1.00\text{e-}01$ , I search the 2 cycles to find the better value.

The below tablet shows the value of 3 best performing value.

lam	val_acc	n_batch	eta_min	eta_max	n_s	n_epochs	
5.478111e-04	0.5266	100	1e-05	0.1	900.0	8	Network 1
1.939191e-05	0.5214	100	1e-05	0.1	900.0	8	Network 2
4.910818e-03	0.5204	100	1e-05	0.1	900.0	8	Network 3

The below chart shows the train and validation loss.

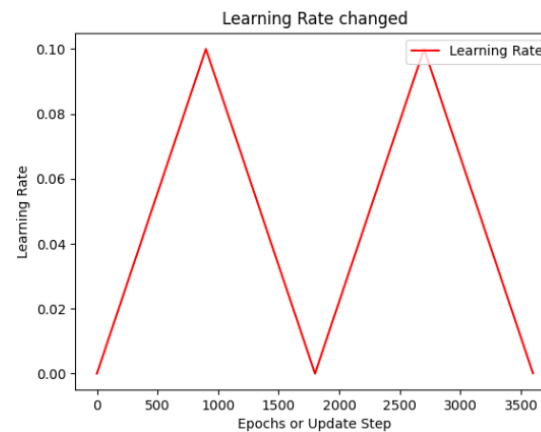
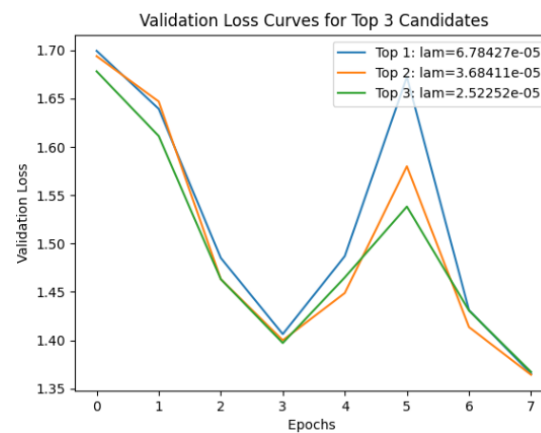
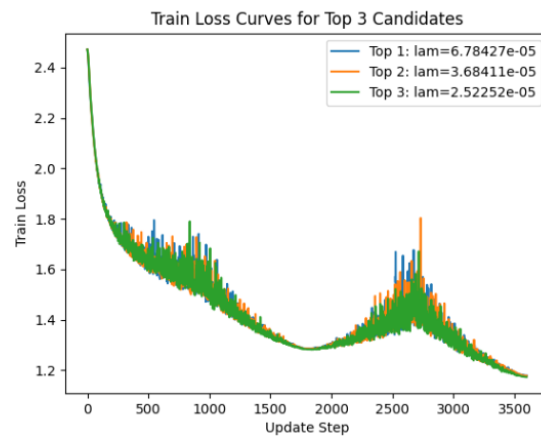


### Question4:

The range of the value I searched during the coarse search for  $\lambda$  is from  $\text{lam\_min} = 10^{-4.7131} = 1.94\text{e-}05$  to  $\text{lam\_max} = 10^{-3.261} = 5.48\text{e-}04$ , I search the 2 cycles to find the better value. The below tablet shows the value of 3 best performing value.

lam	val_acc	n_batch	eta_min	eta_max	n_s	n_epochs	
6.784271e-05	0.5258	100	1e-05	0.1	900.0	8	Network 1
3.684114e-05	0.5256	100	1e-05	0.1	900.0	8	Network 2
2.522519e-05	0.5230	100	1e-05	0.1	900.0	8	Network 3

The below chart shows the train and validation loss.



### Question5:

The final accuracy is 0.5157, the below chart show the loss of train and validation.

