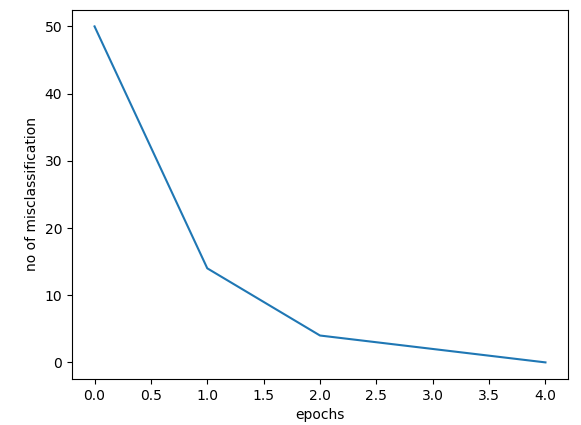
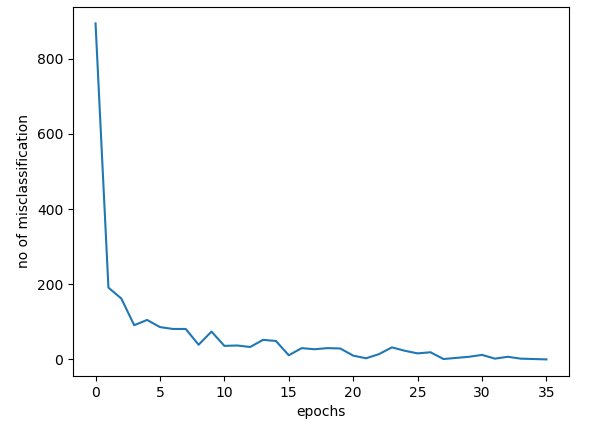
q2) Following are the answers for the report

f) 

Testing error (on all 10000 test images): 4603

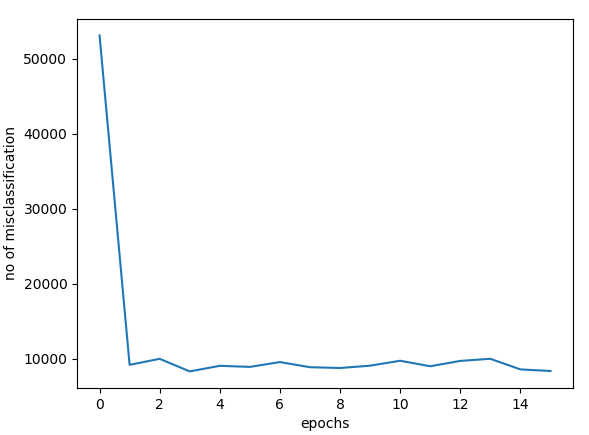
testing error: 0.46 (46%)

So the testing error is higher even when training error is 0 because we have trained only on limited sample(50) so the model can’t accurately predict the testing sample

g) 

testing error = 1794/10000 = 0.1794

So as compared to the previous part it takes more epochs (35 as compared to 3 before) to converge as we have increased the no of training samples in this case. The test error has also dropped as compared to before as more training samples were used to train the neural network, so it could learn the weights more accurately.

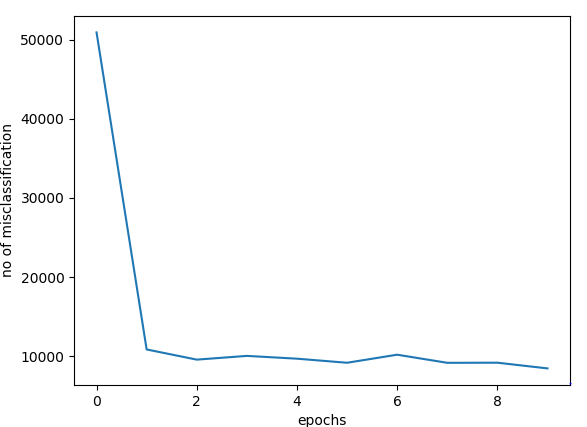
h) 

testing error = 0.1496

So as compared to before the algorithm doesn’t seem to converge, so had to stop after 15 epochs. As the data is not clearly linear separable so the algorithm is not able to converge so it comes close to 0.14 (approx) but it doesn’t seem to converge to zero. The testing error seems to come down as compared to previous part, as we have used all of the training samples.

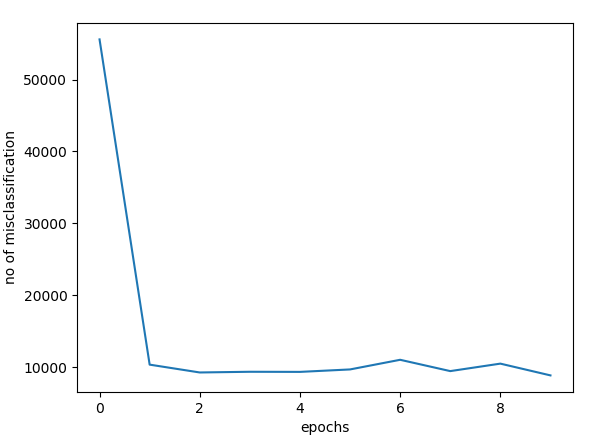
i) Choosing e = 0.15 as it seem to converge along that point

run 1 (n=60000,rate=0.5,e=0.15)



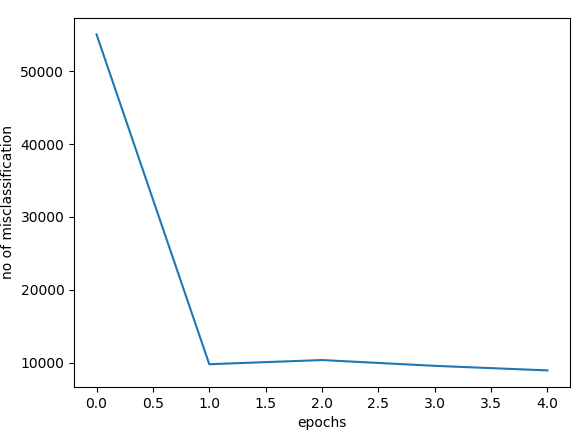
testing error = 0.1677

run 2 (n=60000,rate=0.5,e=0.15)



testing error = 0.1567

run 3 (n=60000,rate=0.5,e=0.15)



testing error = 0.1785

Even with the same value of learnig rate and no of samples, the above runs had different convergence. Different set of initialisation of weights lead to different convergence. Some converge faster and some takes more epochs to converge. The input sample are not clearly linearly separable so the PTA algo is doing its best to get to the minimum error but sometimes after changing weights it might lead to more error and might take longer to converge. Testing error of almost similar but not identical as the three different started with different set of initial weights hence they will have different set of convergence and different W vector, as there is not one correct set of weights but rather its dependent of the initial weight initialisation.