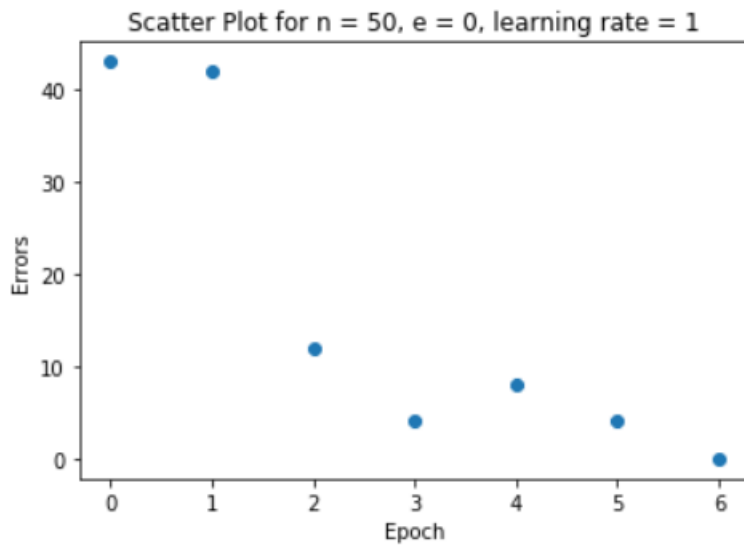
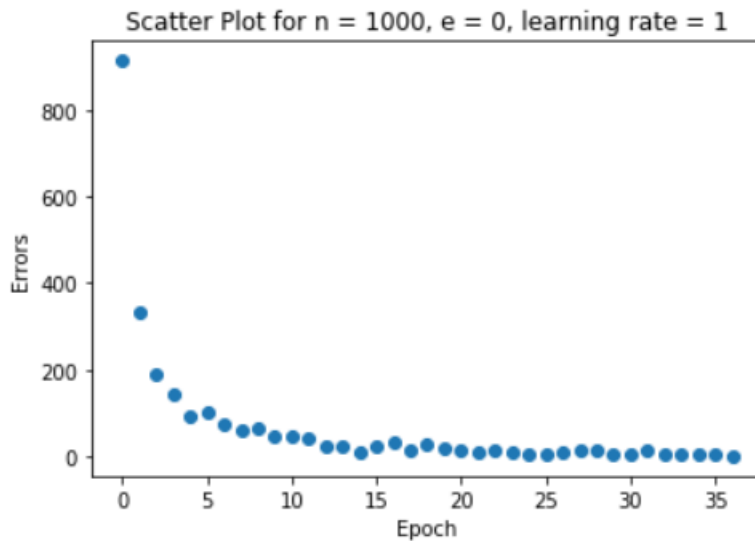


2)f)



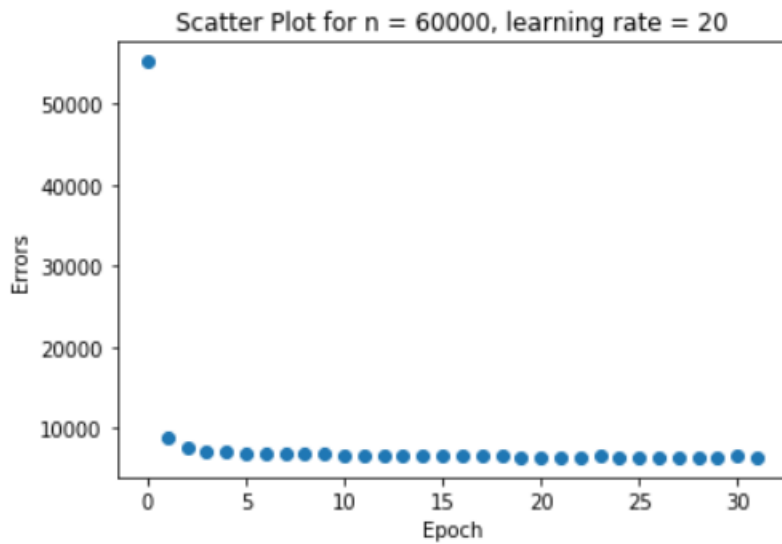
The percentage of testing error for training with $n = 50$, $e = 0$, learning rate = 1 is 43.06. The significant discrepancy between the percentage error is due to a smaller number of samples (50) in training.

g)



Since we used a greater number of samples for training this time, we got less percentage testing error of 17.22. It makes sense as we get more optimal weights with a greater number of training samples.

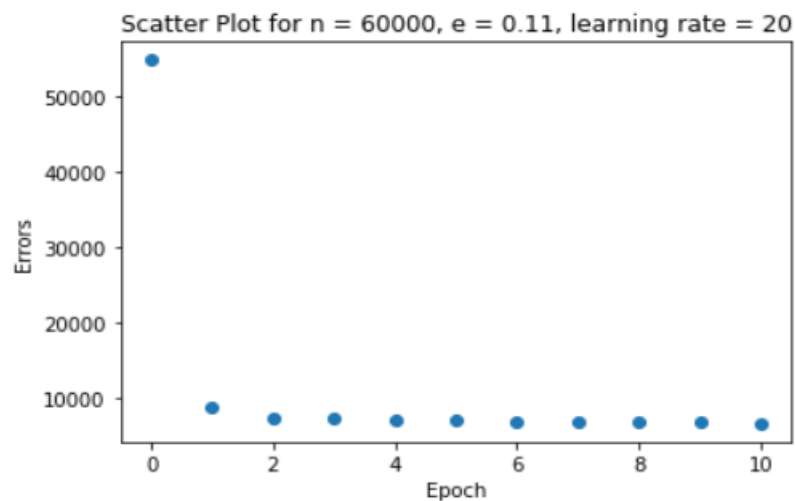
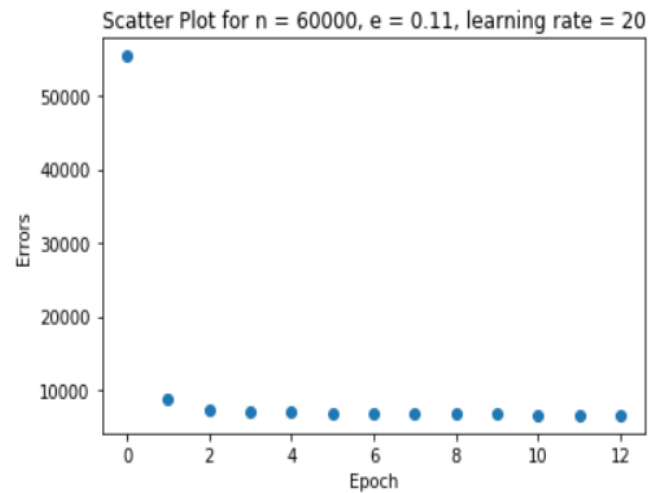
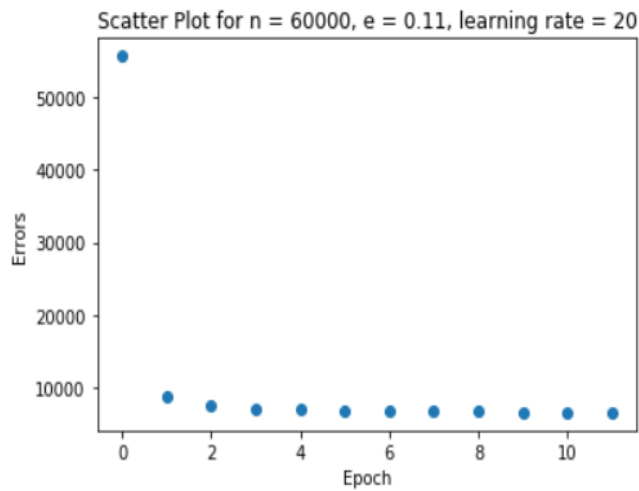
h)



The PTA didn't converge with $n = 60000$, $e = 0$, learning rate = 20. So, I modified the algorithm to stop after 30 epochs. The number of misclassifications reduced drastically in second epoch, but they didn't reduce much from second epoch after each epoch.

In this case, I got lesser percentage testing error of 11.78 because we used 60000 samples for training.

I)



Based on my results, for error threshold of 0.1, the PTA didn't converge (or might have converged after a lot of time) with learning rates: {1,10,20}. And with a greater learning rate, the PTA converged quickly. But the percentage of testing error has no correlation with the learning rate.