REPORT HOMEWORK 3 – PTA

(f) With n = 50, step = 1 and eps = 0 the errors on the test are 4559, due to underfit: the model has few training sample to generalize well to unseen data, as we can also see on the percentage of misclassification the test set that is equal to 45.59%, while for the training, having 0 misclassifications, will be equal to 0%.

(g) With n = 1000, step = 1 and eps = 0 the errors on the test are 1785, the percentage of errors on train is 0%, while on test 17.85%. We can see that we have an improvement in the generalization of the algorithm, and performances are enough good, with an accuracy of about 82%. We have no more underfit.

(h) With n=60000 and eps=0 the algorithm doesn't converge because it doesn't find the right weights in order to have 0 misclassification, this because we have, even if minimum, an error on classification; due also to the fact the feature space is high, with one input for each pixel of the image.

(i) Starting with n=60000 and a eps=0.15 and step=1, the network performed with 1552 errors in the test set, a percentage of misclassification for the test set equal to 15.52%, while for the training will be equal to 13.89%. Is a good result because the percentage of errors on the training is near to the one of the test. The model can for this reason perform well to unseen data.

Choosing different weights is possible to see, from the plots of epochs and misclassification errors that the behavior is different according to the fact the weights could be nearer of farer from the optimal ones. From the point of view of results, they performed in the same in the training because the condition on the while is always the same. For the test set the results are different, depending on how much the model has been trained well. Changing the step from 1 to 10 didn't influenced the training, while with a step of 100, the performances got worse.