PROGRAM DESCRIPTION

The program pick uniformly at random 3 weights, respectively w0,w1 and w2:

- w0= -0.18030709
- w1= 0.26574104
- w2= 0.94162416

So the vector of weights OMEGA in this case will be :

$$\Omega=[-0.18030709,\ 0.26574104,\ 0.94162416]$$

Now we pick n = 100 vectors x1 , . . . , xn independently and uniformly at random on $[-1, 1]^2$ and we call the collection of these vectors S.
Then S1 $\subset$ S will denote the collection of all x = [x1 x2] $\in$ S satisfying [1 x1 x2][w0 w1 w2]^T ≥ 0, while S0 $\subset$ S will denote the collection of all x = [x1 x2] $\in$ S satisfying [1 x1 x2][w0 w1 w2] ^T < 0.

We pick w0′, w1′ and w2′ independently and uniformly at random on [−1, 1]:

- w0′= -0.00269253
- w1′= -0.56700527
- w2′= -0.66693574

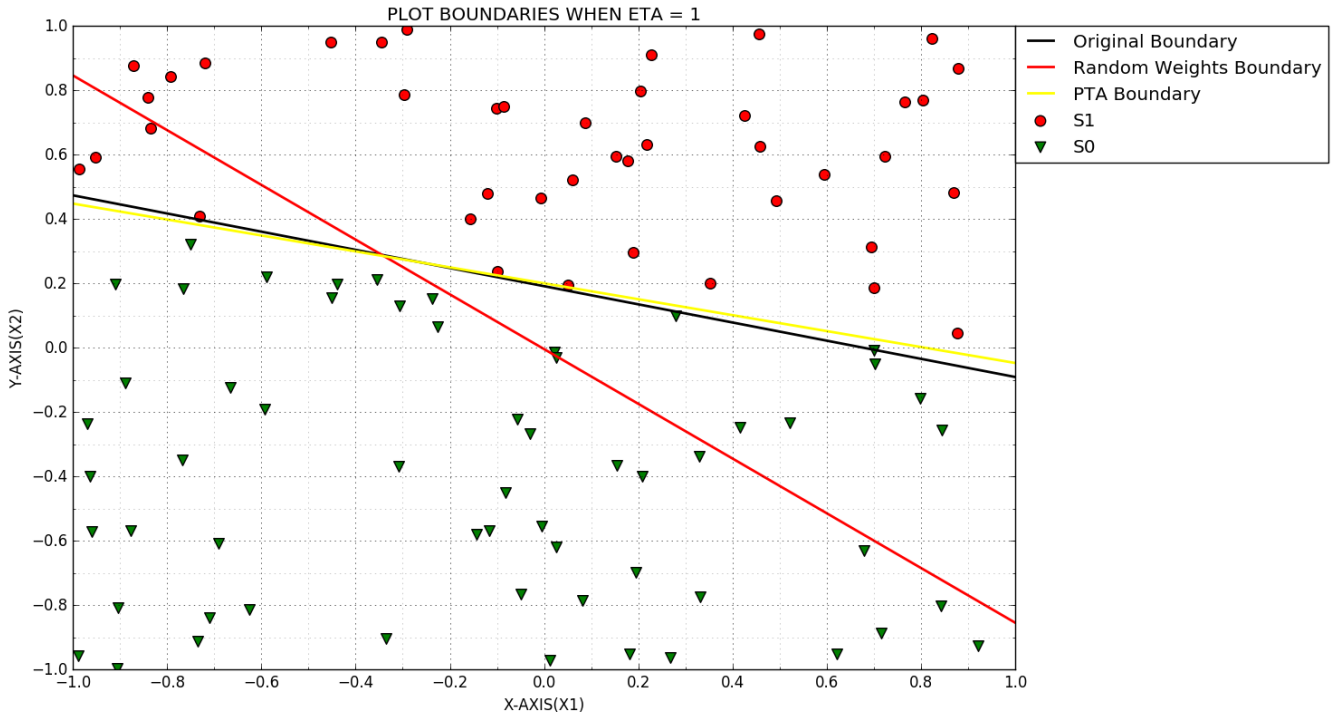
So the vector of weights OMEGA′ in this case will be :

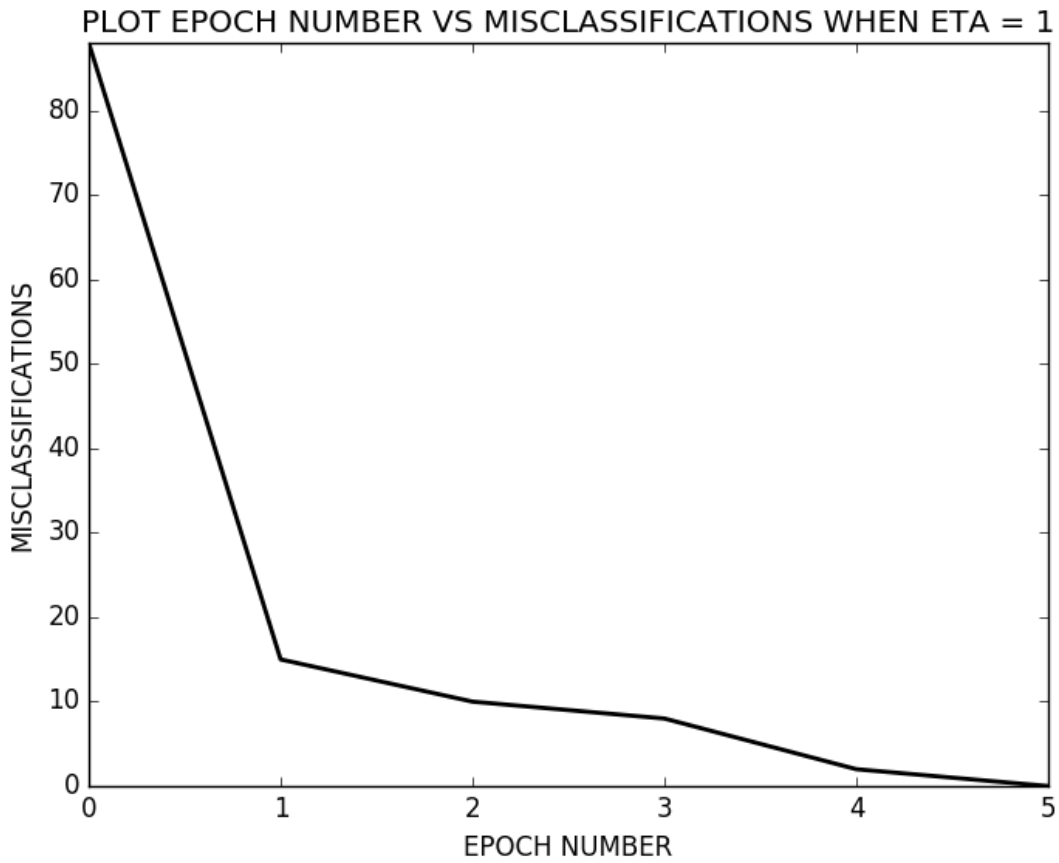
$$\Omega'=[\ -0.00269253\ -0.56700527\ -0.66693574]$$

Now we will apply the PTA using n=100 and different parameters η.
While η will be changed, Ω, Ω' and S will remain the same.

PTA with n=100 and η=1 :

- Total number of epoch=5
- Final Ω=[-1.00269253, 1.23861838, 5.0024749]

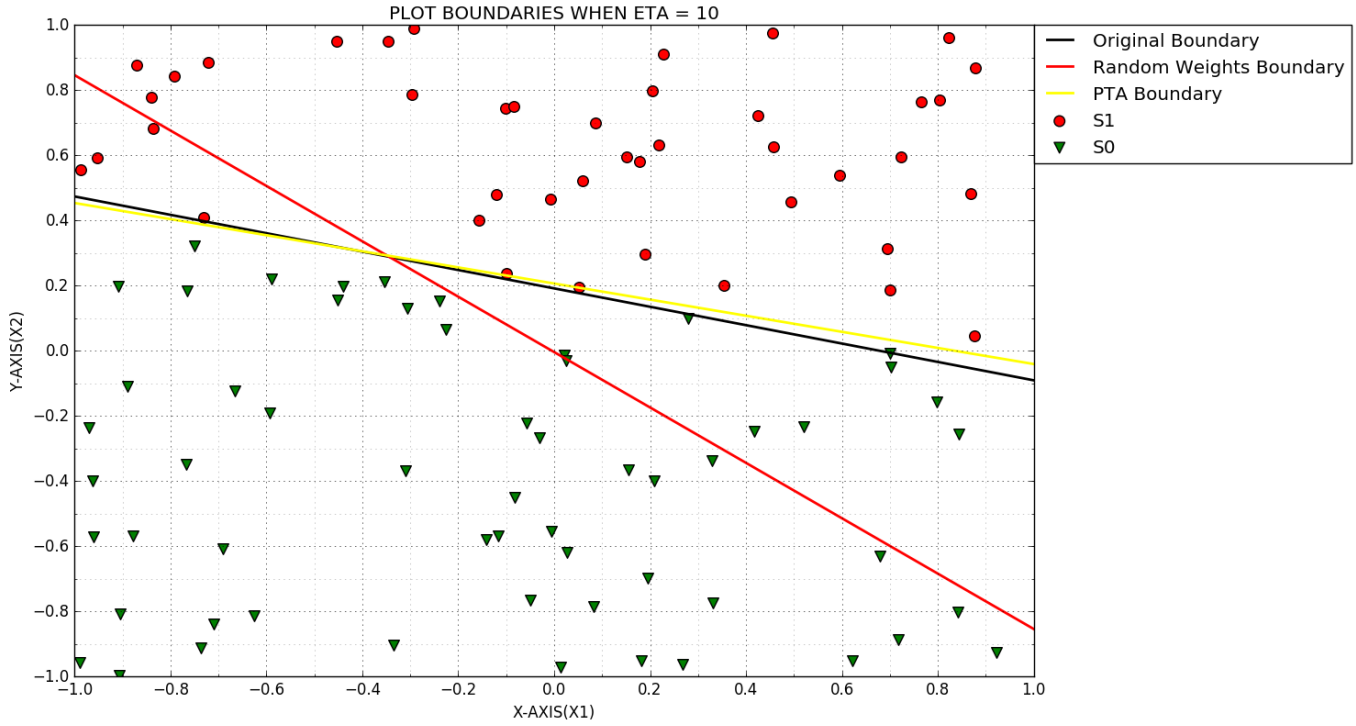


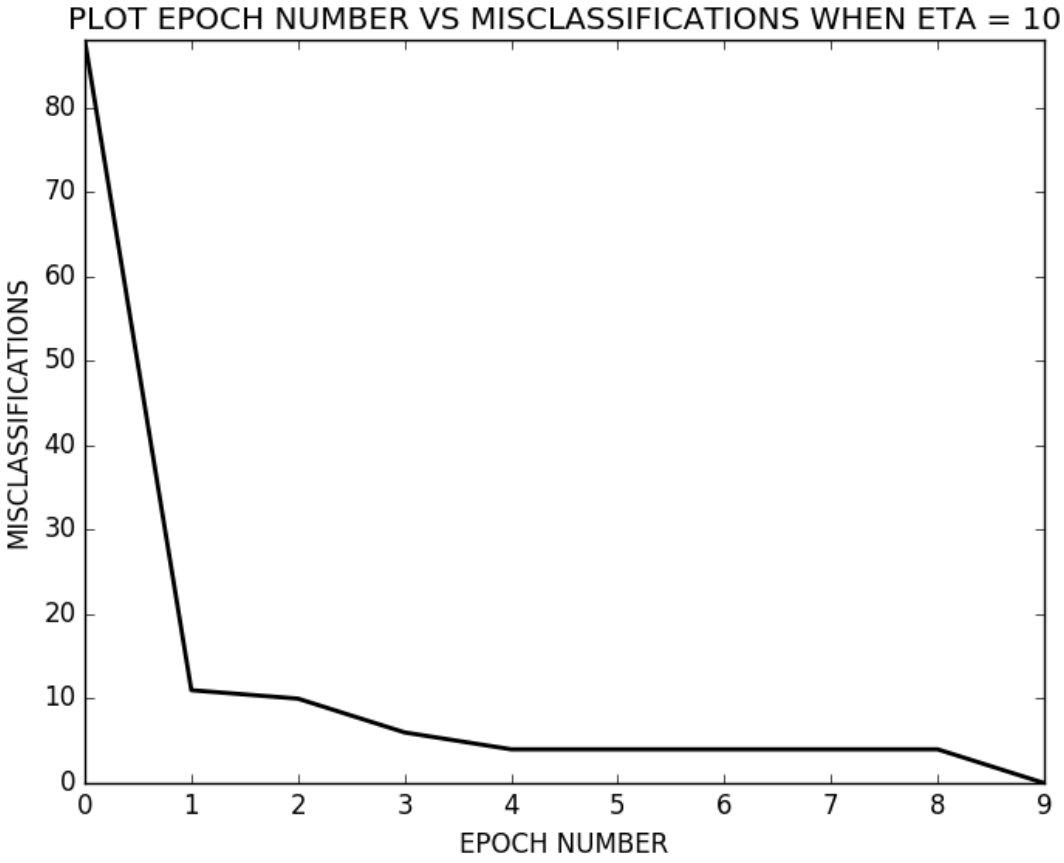


PTA with n=100 and η=10 :

- Total number of epoch=9

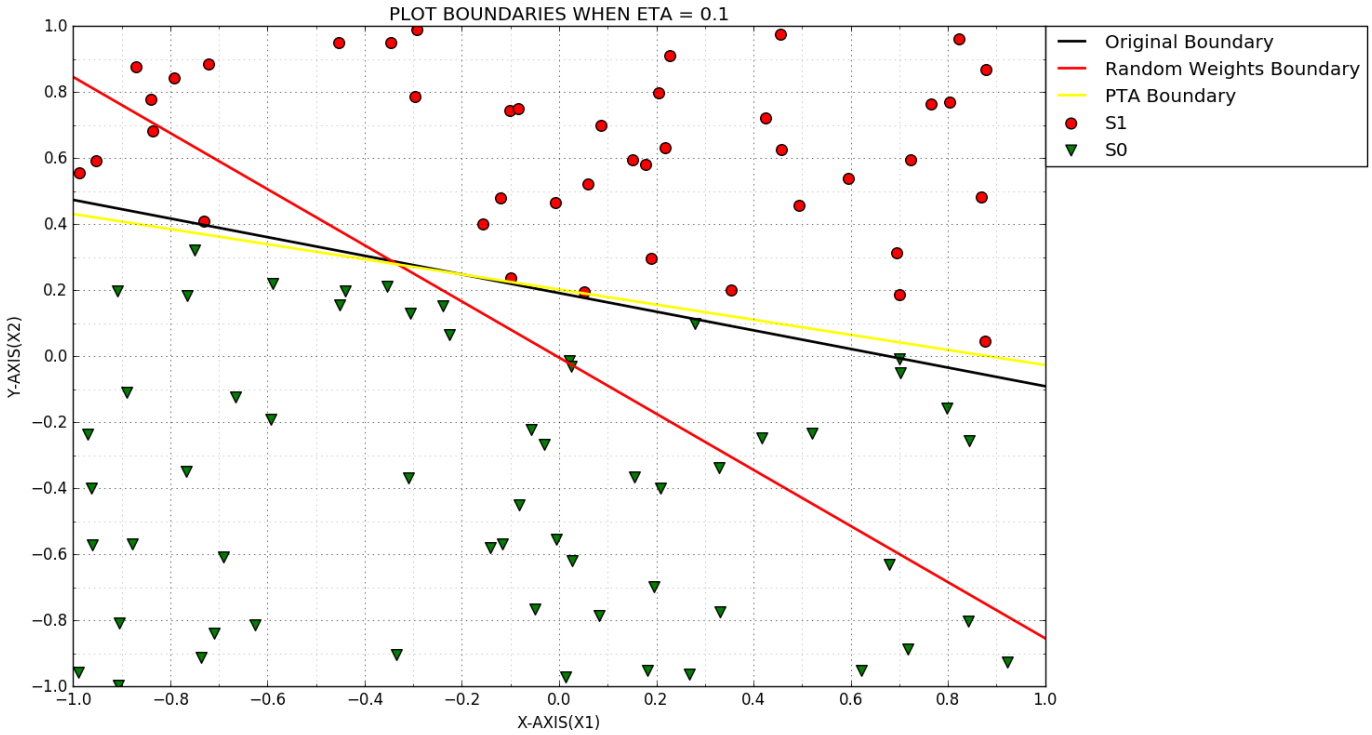
- Final Ω=[-10.00269253 11.977707   48.45247227]

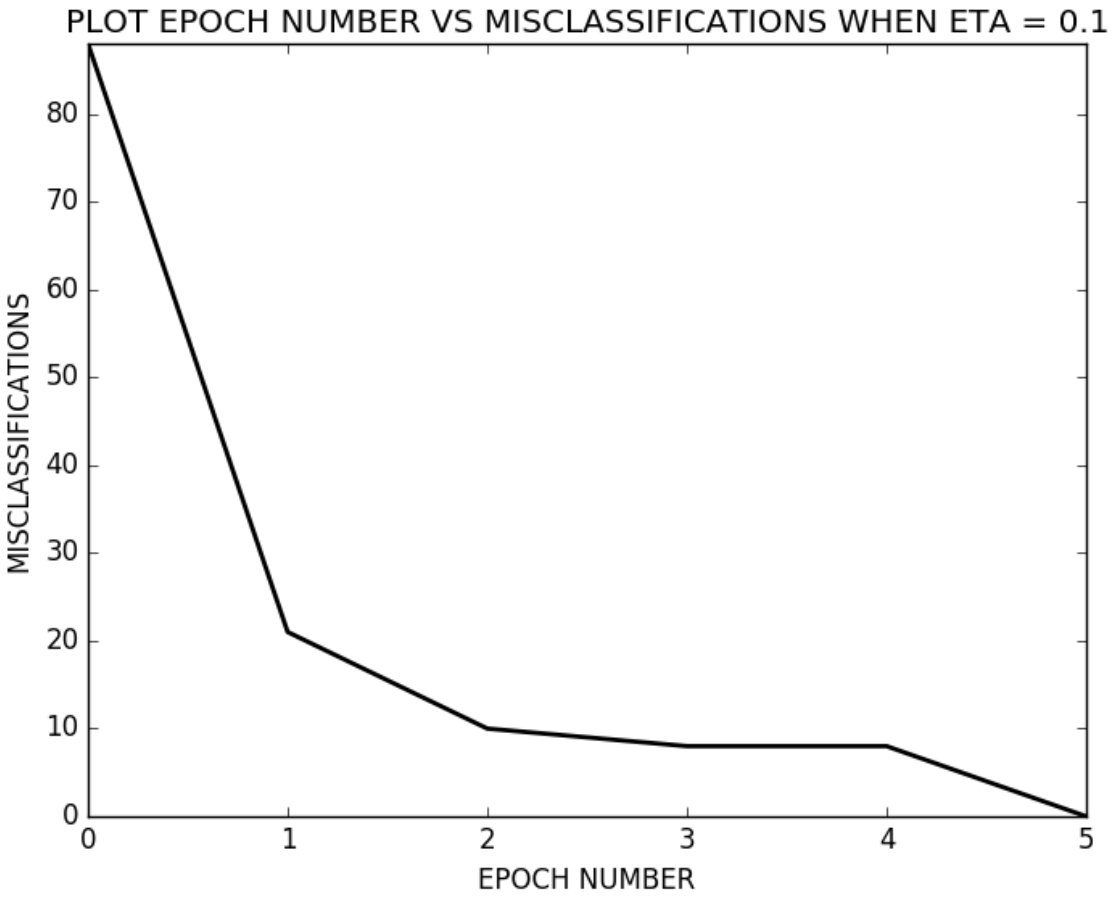




PTA with n=100 and η=0.1 :

- Total number of epoch=5
- Final Ω=[-0.10269253 0.11617793 0.50800075]





Analyzing the final weights obtained with the different values of η, compared with the "optimal" ones, we can notice that when η is big(equal to 10) the weights are very large, while when η is equal to 1, they are smaller compared to the previous one but always larger than the "optimal" . Finally, when η is equal to 0.1 we can notice that the value of the weights is much smaller and closer to the "optimal" weights.
Looking at the graph we can see that anyway even if η changes, the final line(yellow line) it's very close to the initial one(black line) for all the values of η. The one that is closer to the original one is the line when η is equal to 0.1(in fact the weights are more similar).
After repeted experiments, we can also say that the changes in the value of η doesn't influence in a significative way the number of epochs. In fact the number of epochs change independently from η, if the random weights remain the same . For this reason we can say that η and the number of epoch are not correlated.

We can now repeat the previous experiments changing the number of vectors from 100 to 1000.
The program pick uniformly at random 3 weights, respectively w0,w1 and w2.
So the vector of weights OMEGA in this case will be :

$$\Omega = [\ -0.09126036,\ 0.1036276,\ -0.15175198]$$

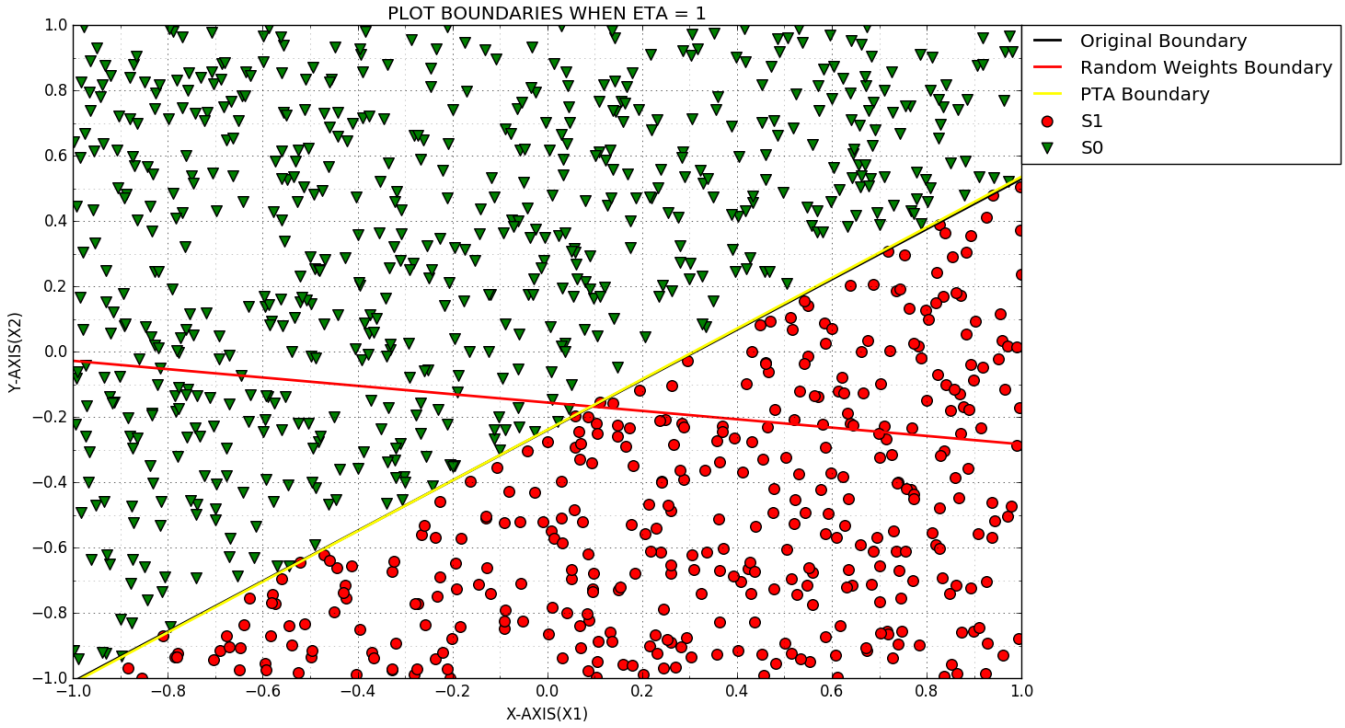
We repeat the same operations as before and pick $w0'$, $w1'$ and $w2'$ independently and uniformly at random . The new vector $OMEGA'$ will be:

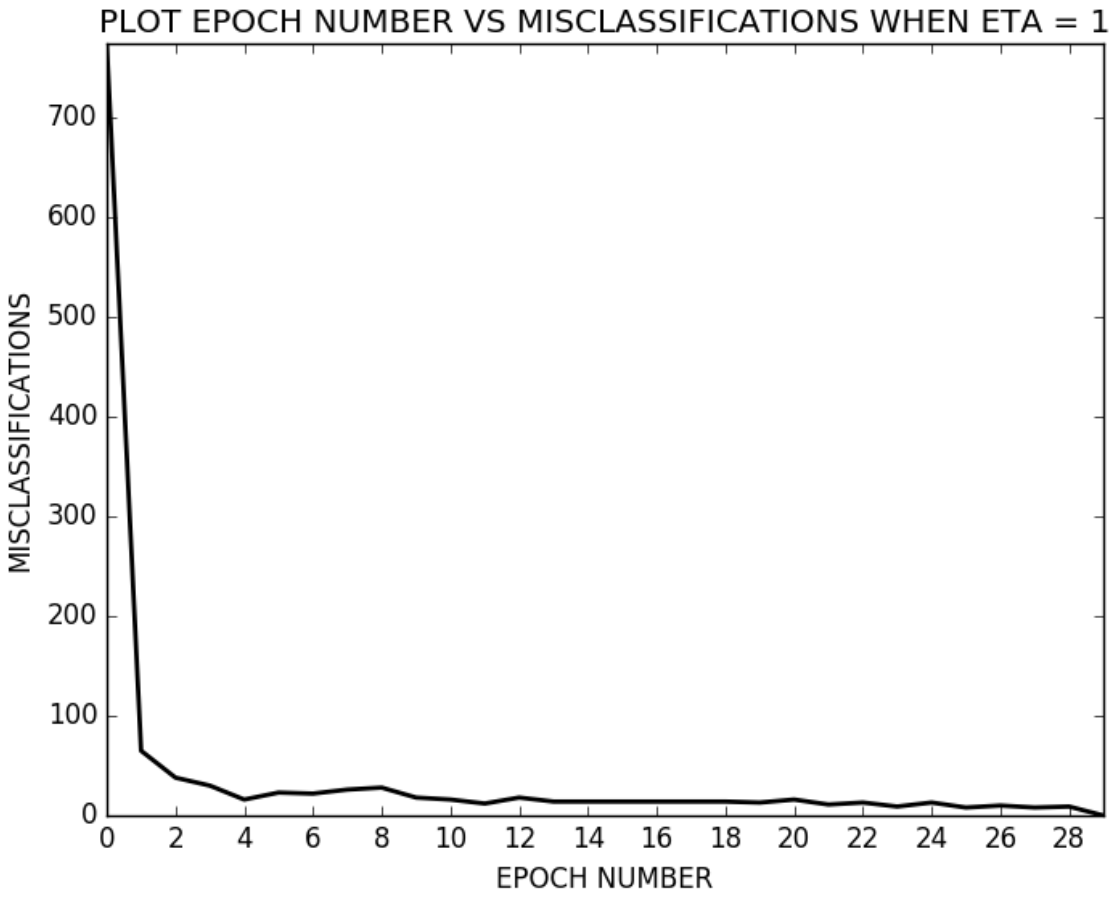
$$\Omega' = [0.5674571,\ -0.19089478,\ 0.10175856\ ]$$

Now we will apply the PTA using n=100 and different parameters η. While η will be changed Ω, Ω′ and S will remain the same.

PTA with n=1000 and η=1 :

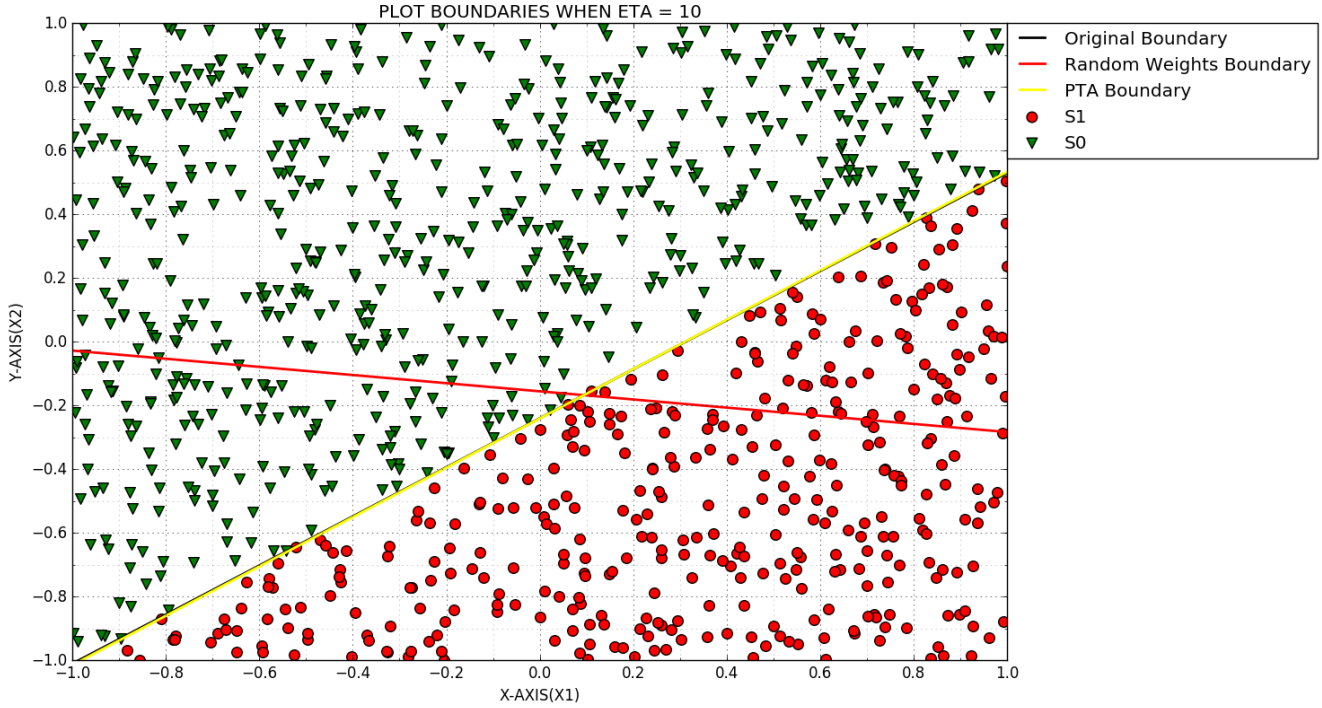
- Total number of epoch=29
- Final Ω=[-3.85173063, 12.48935967, -16.11073203]

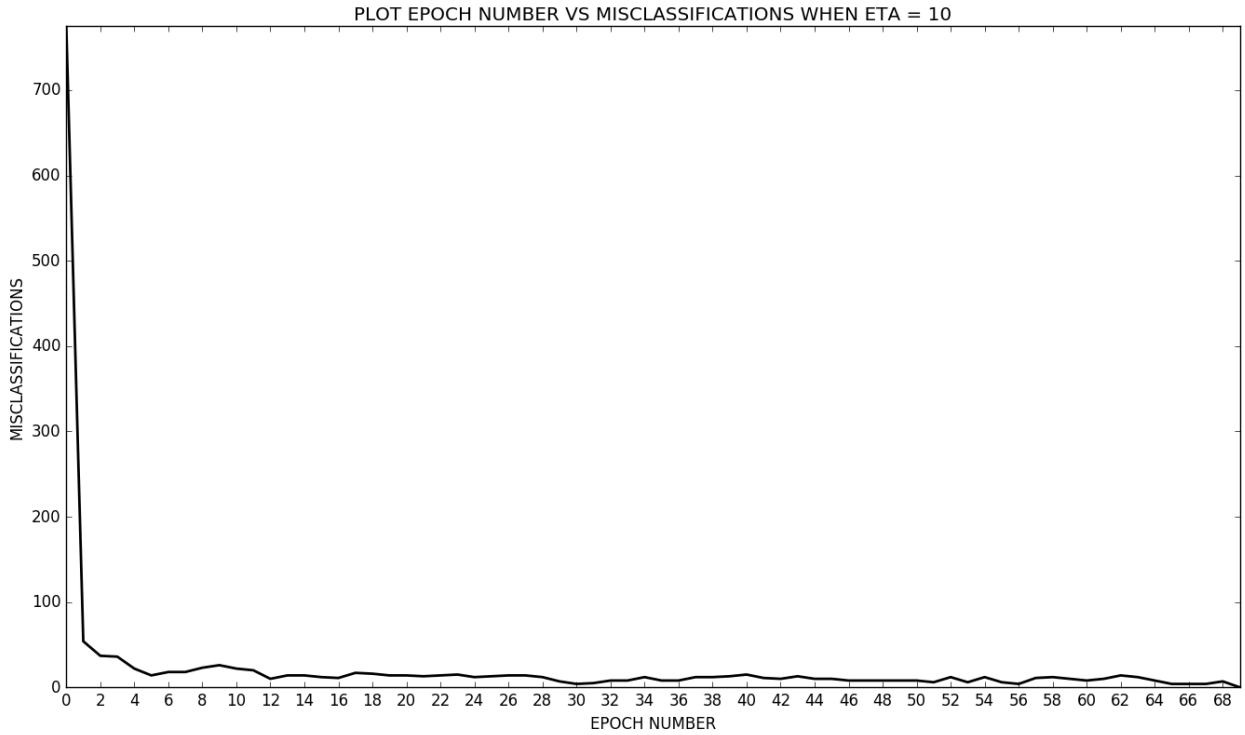




PTA with n=1000 and η=10 :

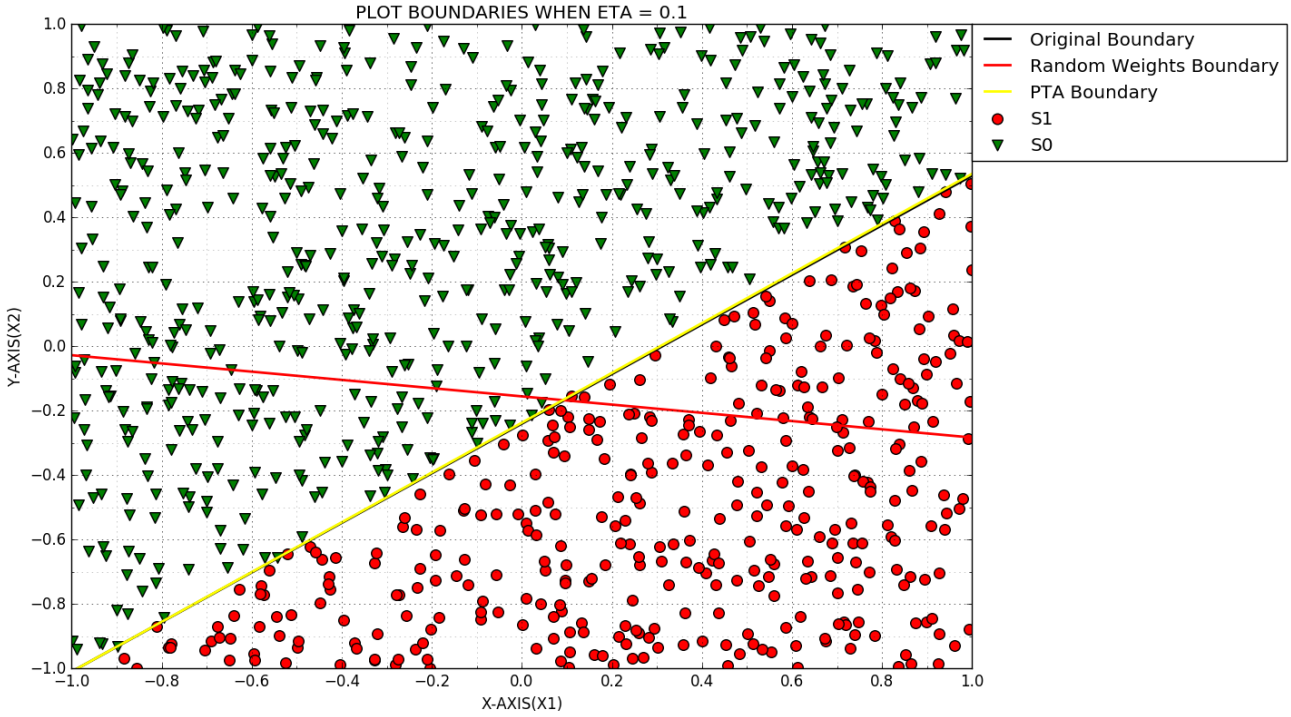
- Total number of epoch=69
- Final Ω=[-49.85173063, 160.28294821, -207.11082377]

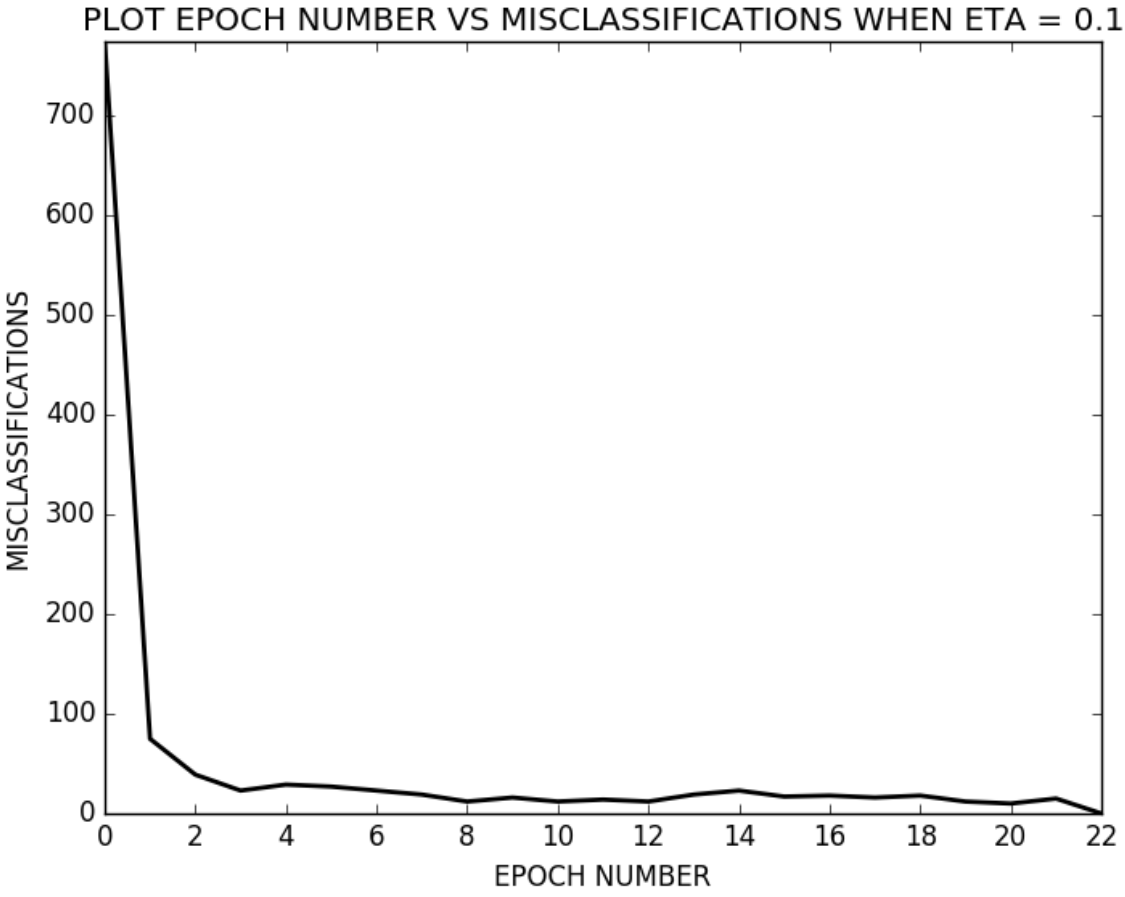




PTA with n=1000 and η=0.1 :

- Total number of epoch=22
- Final Ω=[-0.35173063, 1.14630675, -1.48311263]





If we had started with different w0 , w1 , w2 , S, w0’ , w1’ , w2’, the effects in terms of η will be the same only on the weights dimensions that will be greater with greater η and smaller with a smaller η. By the way, we expect that the number of epochs won’t change in the same way, because as we said before they are not correlated.

The main difference using n=1000 samples is that the weights are much bigger,and usually also the number of epoch is greater, even if this is not always true and more experiments will be necessary. If we change η we obtain exactly the same effect as before,in fact η and the number of epoch are not linearly correlated.