Tuesday, September 19, 2023 7:44 PM

Q1.

I used a seed value of 555 to ensure same results across different runs, although, I have run the program without a seed to determine the trends while answering questions.

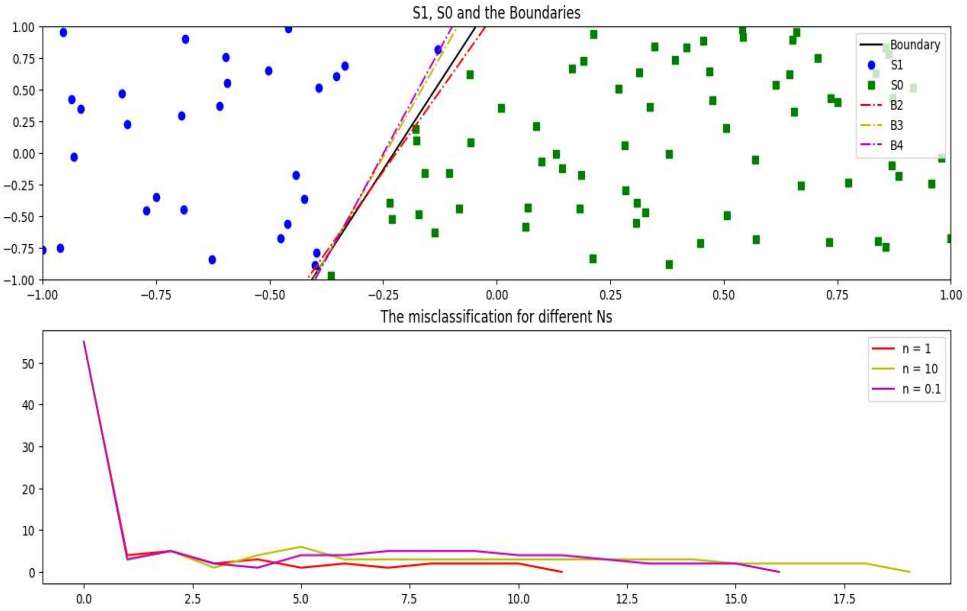
a, b, c:
w0, w1, w2 picked uniformly are: -0.15332737462764612 , -0.6778882424805346 , 0.1224450872652012

d, e, f:
too long to fit in this report, please uncomment the lines 38-40 and run the program to see the results

g:

There are two subplots in the picture below. The first one shows the 100 vectors generated and the boundary (names Boundary in Black) plotted using the weights w0, w1, w2. The other boundaries B2, B3 and B4 are the boundaries produced by the PTA which I will go over in the next parts.

The blue circles represent the S1 and green squares represent the S0



h:

The random weights picked for PTA are:

listed as (w0', w1', w2')
[-0.38782470612811104 , 0.8306490357910934 , -0.21339803637345067]

The weights produced by the PTA and the total epochs taken are below:

total epochs taken for n = 1 : 11
the final weights for n = 1 : [-1.38782471 -6.27615019 1.23643475]

the weights are very different when compared to the original weights
but they produce a line (wt0 + wt1*x1 + wt2*x2) that has slope and intercept close to the original line/weights.

for w0,w1,w2
slope = -w1/w2 = -(-0.6778882424805346/0.1224450872652012) = 5.53626329664261
intercept = -w0/w2 = -(-0.15332737462764612/0.1224450872652012) = 1.252213363983788

for wt0, wt1, wt2 (weights produced by PTA for n = 1)
slope = -wt1/wt2 = -(-6.27615019/1.23643475) = 5.07600598
intercept = -wt0/wt2 = -(-1.38782471/1.23643475) = 1.12244072

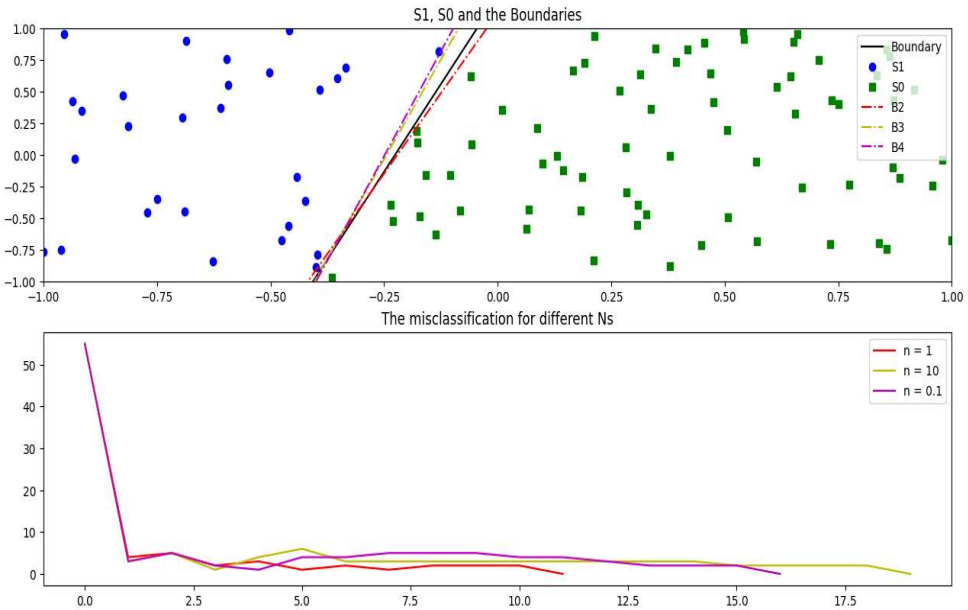
produced by PTA vs optimal/original
1.12244072 vs 1.252213363983788
5.07600598 vs 5.53626329664261

And this is due to the fact that lines in only a certain area between the clusters can act as a boundary.

i:

Using the same plot as above, the second subplot shows the graph of epoch number vs the number of misclassifications for each epoch.

The line for n = 1 is the red line. We can see a sudden drop just after the first iteration.



j, k:

The lines for n = 10 and n = 0.1 are in the same subplot with n = 1.

n = 10 has yellow color, n = 0.1 has magenta color for both misclassifications and boundaries subplots.

total epochs taken for n = 10 : 19
the final weights for n = 10 : [-20.38782471 -83.27835992 13.22169833]

total epochs taken for n = 0.1 : 16
the final weights for n = 0.1 : [-0.18782471 -0.7550908   0.1141193]

l:

after testing using several seed values, I saw a trend that more often than not, higher learning rate meant more epochs are required to produce weights that converge. The desired range of slope and intercept we are trying to reach here have a numerator and denominator each (numerators: w0, w1 and denominator: w2), and higher learning rate means that a higher number is added/subtracted from the starting weights since we use the following formula:

w = w - n*x_i or w = w + n*x_i

Manipulating using higher numbers from numerators and denominator takes more iterations to adjust towards the desired output whereas lower numbers can do this quickly, but going too low like 0.1 can also take a longer time to adjust the weights to produce the desired ratios.

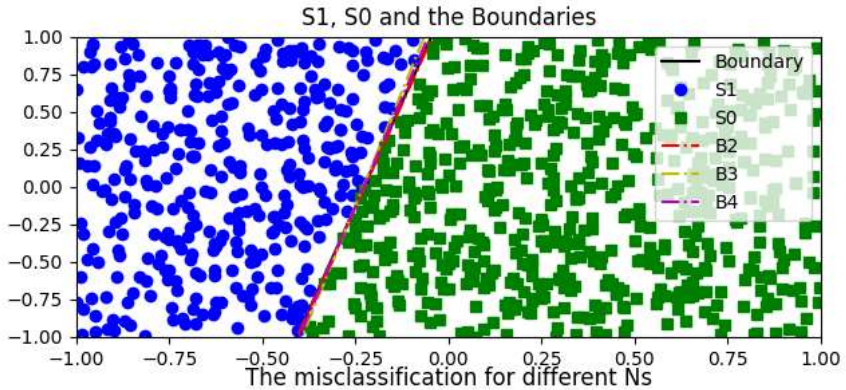
We also see misclassifications drop at first and rise a bit to fall again, as so on until convergence since we are adjusting the weights slightly each iteration to move towards the desired range for each weight.

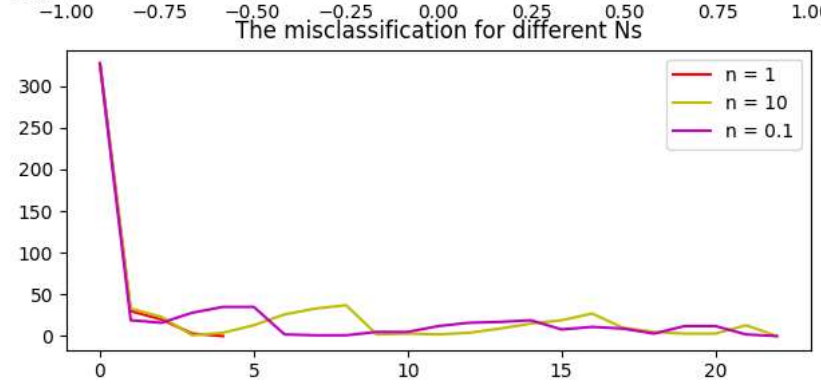
m:

after testing using several seed values, thus producing different weights each time, I observed similar results and effects across different learning rates, where larger and lower numbers (like n = 10 and n = 0.1) taker more epochs to produce desired results than a middle range of numbers (n = 1).

n:

graph for N = 1000





w0, w1, w2 picked uniformly are: -0.15332737462764612 , -0.6778882424805346 , 0.1224450872652012

The random weights picked for PTA
are: -0.9559892321579118 , -0.8294711453722292 , -0.6054554642219308

total epochs taken for n = 1 : 4
the final weights for n = 1 : [-1.95598923 -8.70831063 1.58143828]

total epochs taken for n = 10 : 22
the final weights for n = 10 : [-40.95598923 -178.87342972 29.28802949]

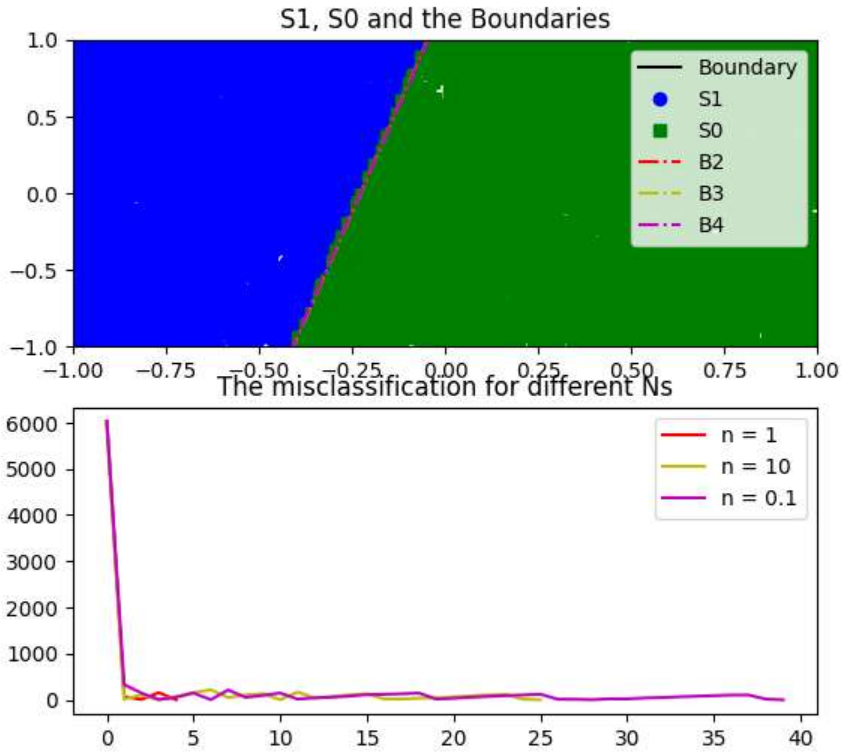
total epochs taken for n = 0.1 : 22
the final weights for n = 0.1 : [-0.45598923 -2.00684476 0.3465435]

We see the similar effect where a n = 1 takes lesser iterations that n=0.1 and n=10.

Although the initial number of misclassifications are higher as compared to N=100 as the initial random weights will misclassify more points than before.

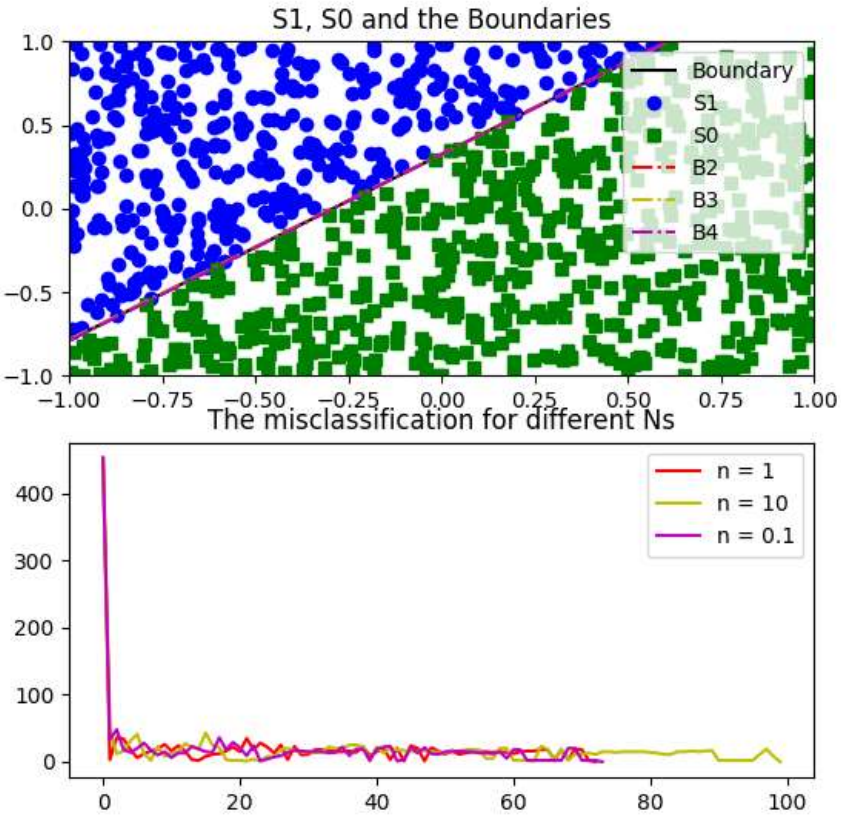
But, we see that now, less number of iterations/epochs are taken for n = 1 to converge and more iterations are taken by n = 10 and n = 0.1. This may be due to the higher sample size, having more points means more numbers are pushing/modifying the weights towards the right direction. But, we converge slower for n = 10 and n = 0.1 as now we are either pushing too much or pushing too less towards the desired range of weights to produce the correct slope and intercept. This is confirmed by N=10000 where n = 1 takes 4 iterations (compared to 4 and 11 before), n=10 takes 25 (compared to 22 and 19

before) and n=0.1 takes 39 iterations (compared to 22 and 16 before), results included below.

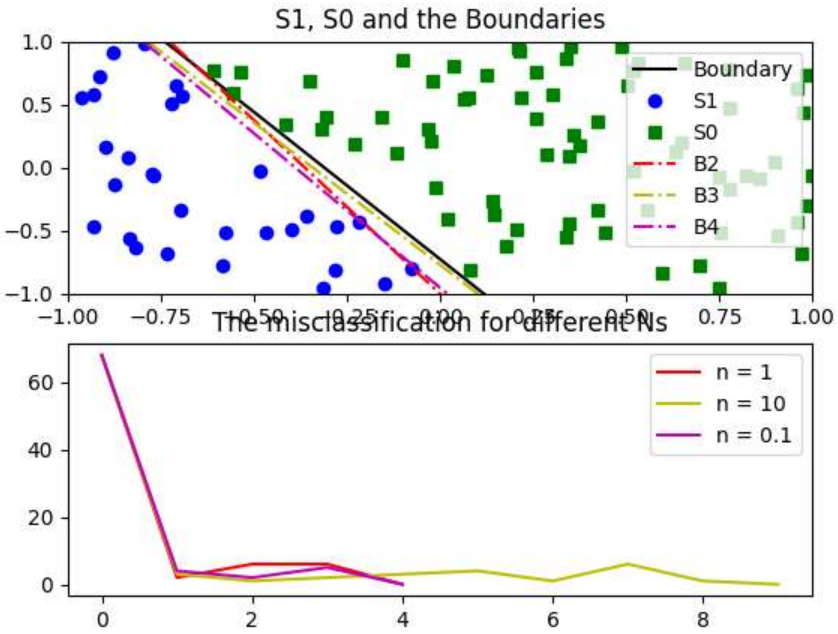


total epochs taken for n = 1 : 4
the final weights for n = 1 : [-4.31048071 -19.00377024 3.42409751]
total epochs taken for n = 10 : 25
the final weights for n = 10 : [-79.31048071 -350.07436604 63.37027861]
total epochs taken for n = 0.1 : 39
the final weights for n = 0.1 : [-0.91048071 -4.01580883 0.73171]

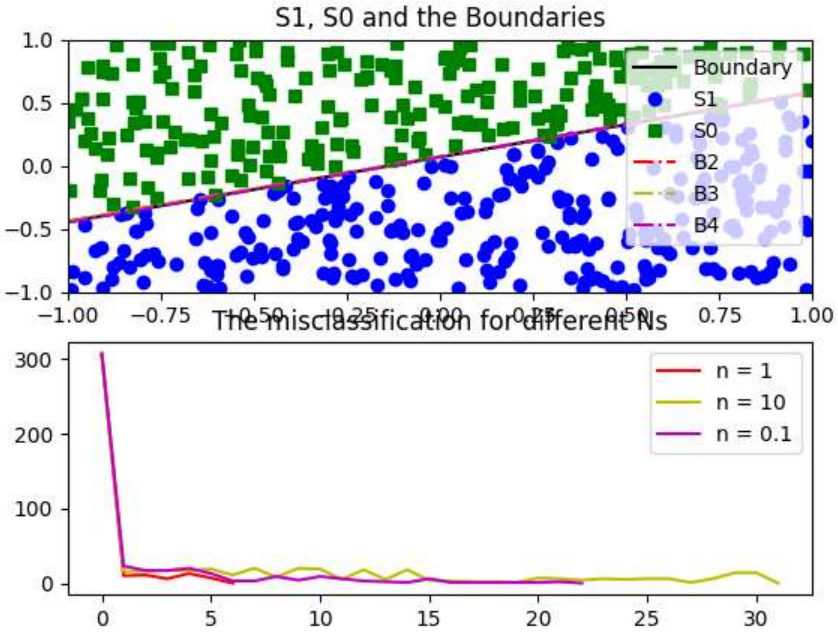
To support the above answers for part (n), (m), (l) I have included the graphs and results for different seed values and N below



w0, w1, w2 picked uniformly are: -0.2185270739572408 , -0.7470280984012727 , 0.6714994102866707
The random weights picked for PTA are: -0.5396910526640928 , 0.5391732596812875 , 0.49201272221460757
total epochs taken for n = 1 : 72
the final weights for n = 1 : [-6.53969105 -22.22786133 20.01641216]
total epochs taken for n = 10 : 99
the final weights for n = 10 : [-70.53969105 -240.1582666 216.2617705]
total epochs taken for n = 0.1 : 73
the final weights for n = 0.1 : [-0.63969105 -2.19482958 1.96807196]

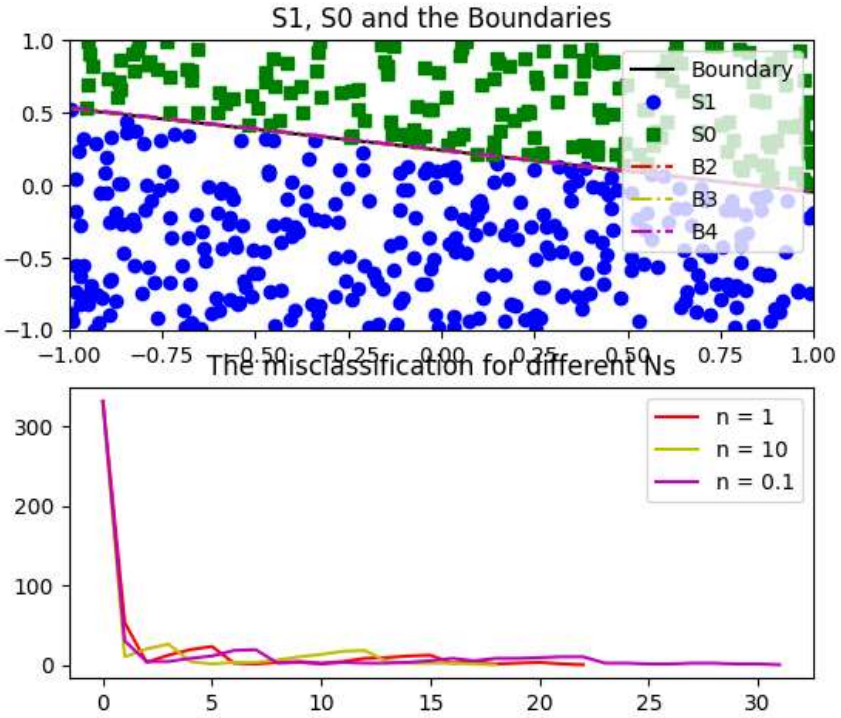


w0, w1, w2 picked uniformly
are: -0.15832253444276284 , -0.5077944250248094 , -0.21754080711107715
The random weights picked for PTA are: 0.27362552661350126 , -0.021275725910583798 , 0.9785381533983888
total epochs taken for n = 1 : 4
the final weights for n = 1 : [-1.72637447 -4.7428845 -1.7226849]
total epochs taken for n = 10 : 9
the final weights for n = 10 : [-19.72637447 -57.51442045 -25.49728494]
total epochs taken for n = 0.1 : 4
the final weights for n = 0.1 : [-0.22637447 -0.58072038 -0.23715402]



w0, w1, w2 picked uniformly are: 0.05052376532245578 , 0.3779780475622132 , -0.7341111112286947
The random weights picked for PTA are: 0.3795285562109463 , 0.01345148092740156 ,

0.5543677290053977
total epochs taken for n = 1 : 6
the final weights for n = 1 : [0.37952856 2.74062174 -5.40136157]
total epochs taken for n = 10 : 31
the final weights for n = 10 : [10.37952856 73.15535425 -143.05799748]
total epochs taken for n = 0.1 : 22
the final weights for n = 0.1 : [0.07952856 0.59801504 -1.15431824]



w0, w1, w2 picked uniformly are:
0.2226171794959747 , -0.26634338734728846 , -0.9227487666378626
The random weights picked for PTA are: -0.7427603617153218 , 0.5323614455802117 , 0.44615121014245784
total epochs taken for n = 1 : 22
the final weights for n = 1 : [3.25723964 -3.86005134 -13.35041562]
total epochs taken for n = 10 : 18
the final weights for n = 10 : [29.25723964 -35.12833293 -120.36877807]
total epochs taken for n = 0.1 : 31
the final weights for n = 0.1 : [0.35723964 -0.428607 -1.46391034]