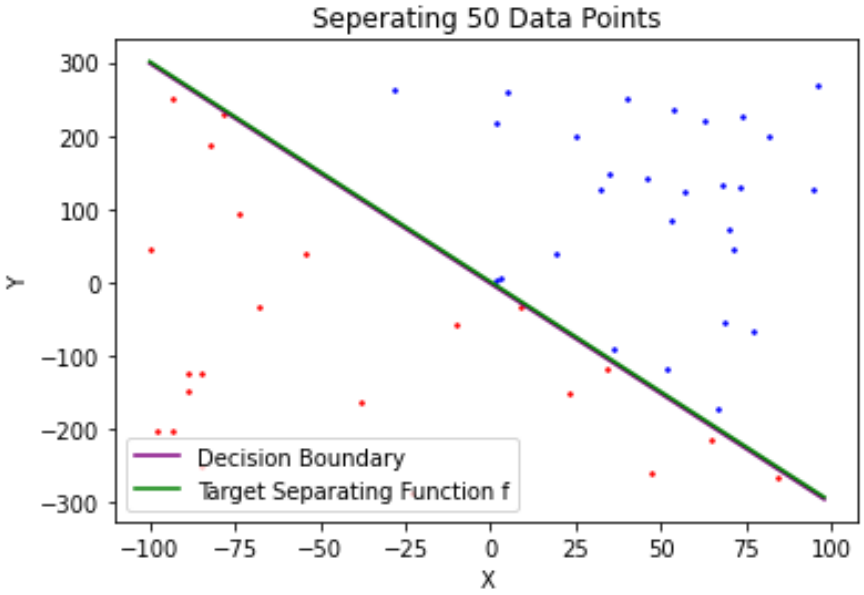
Learning rate is 0.01 for all steps.

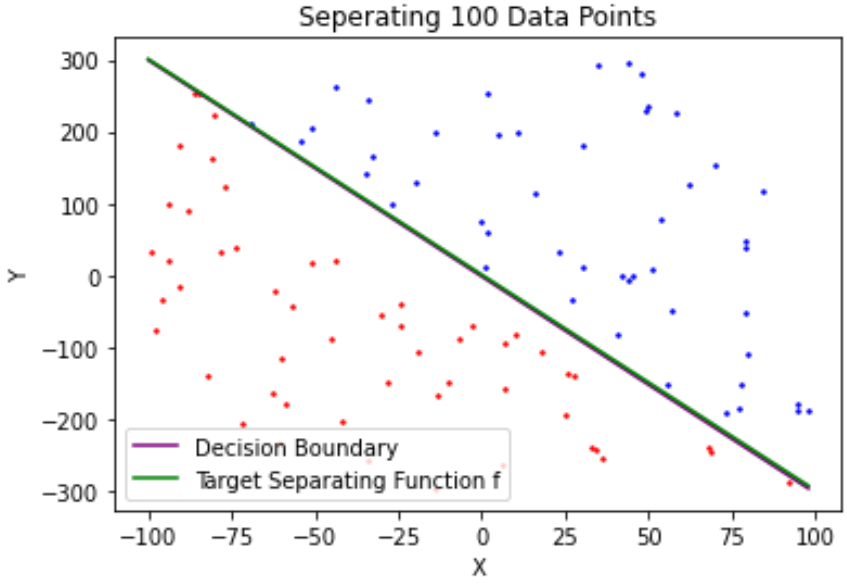
Step 1

Weight vector W: [0.72 3.11 0.86]
Data Count: 50
İteration Count: 10000



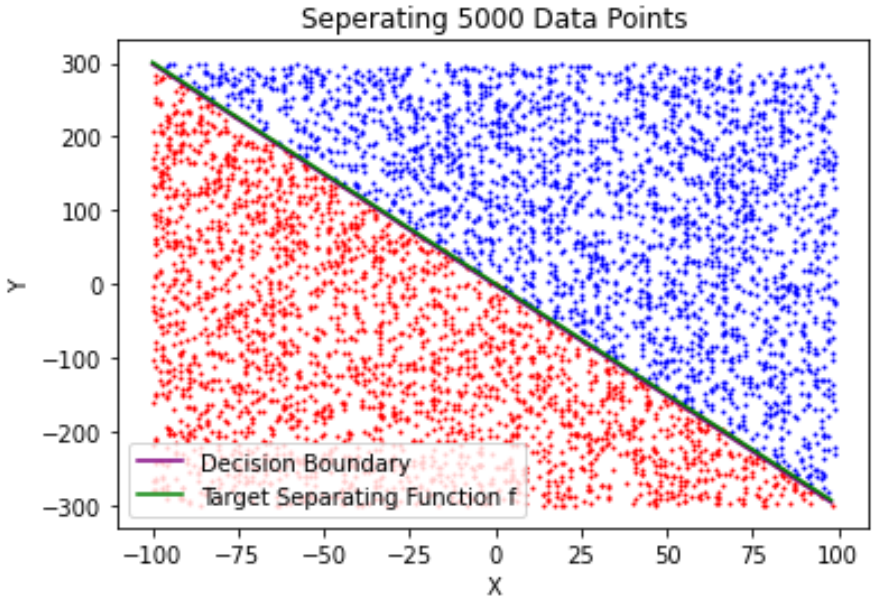
Step 2

Weight vector W: [-12.21 24.1   8.15]
Data Count: 100
İteration Count: 1000



Step 3

Weight vector W: [-16.59 27.65   9.21]
Data Count: 5000
İteration Count: 200



As the number of data points increases, target separating function f and calculated decision boundary overlap better. When data is sparse, the distance between the points with different labels is greater which means at the end, the decision boundary can be located anywhere between the interval. The larger the number of points, the narrower the distance, so the decision boundary gets closer to the target function.

Iteration number decreases as the number of data points increases. The weight vector updated for each mismatching point label at each iteration. As the number of data points increases, the number of updates of the weight vector increases, so we can reach the desired values in less iterations.