ID: 170104082

a) In task 2, why do we need to take the sample points to a high dimension?
Ans: We need to take the sample points to a high dimension if we can't correctly classify the data points in lower dimensions. This φ function does the conversion from lower dimension to higher dimension.

x1w1+x2w2+w0=0   ==> linear equation

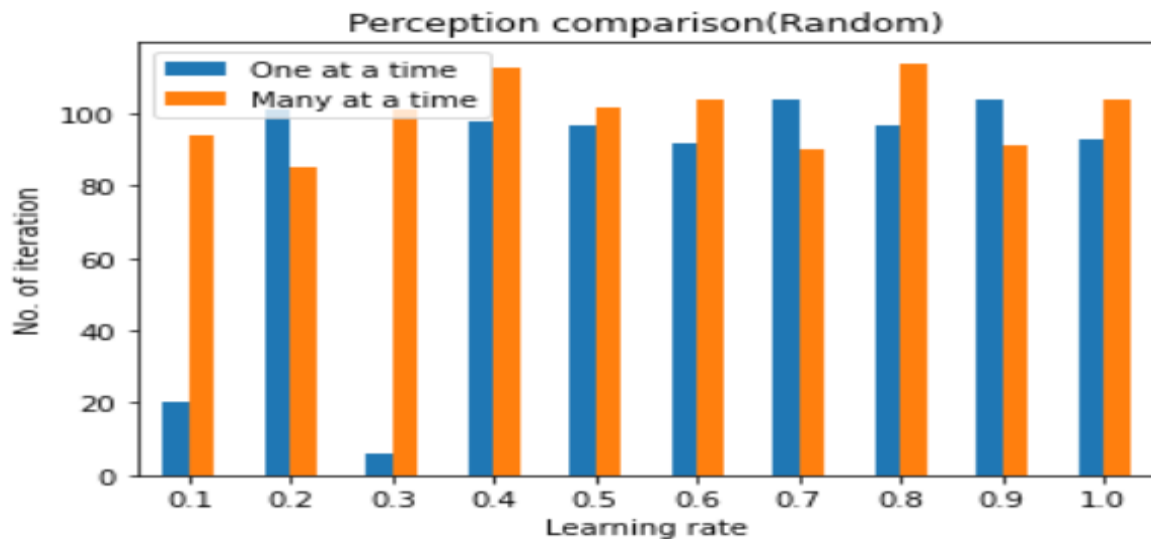$x1^2w1 + x2^2w2 + w0 = 0$ ⇒ Nonlinear equation

Fi
Function

$$(2,4) \quad \Rightarrow \quad \begin{pmatrix} x1^2 \\ x2^2 \\ x1^*x2 \\ X1 \\ x2 \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} 4 \\ 16 \\ 8 \\ 2 \\ 4 \end{pmatrix}$$
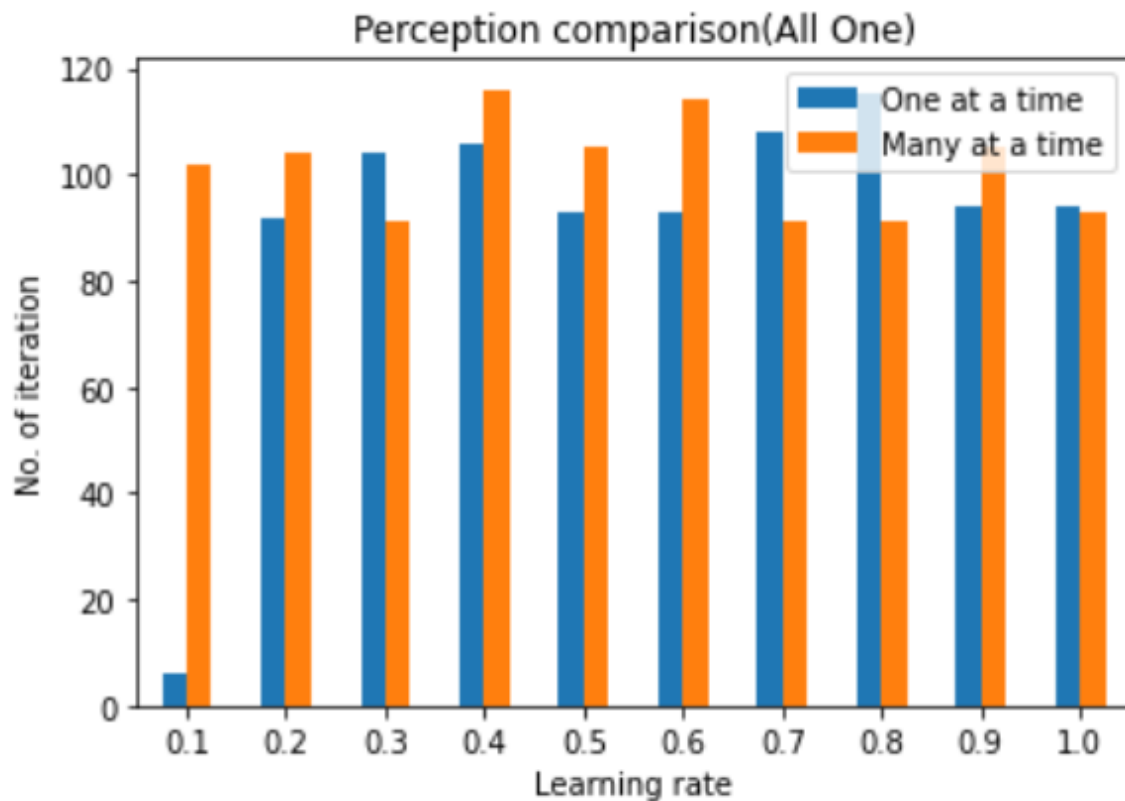
2D

5D

b) In each of the three initial weight cases and for each learning rate, how many
updates does the algorithm take before converging?

Ans:

Case-1:(When weights are initialized randomly)

Case-2:(When all the weights are initialized as 1)

Case-3:(When all the weights are initialized as 0)

Perception comparison(All zero)