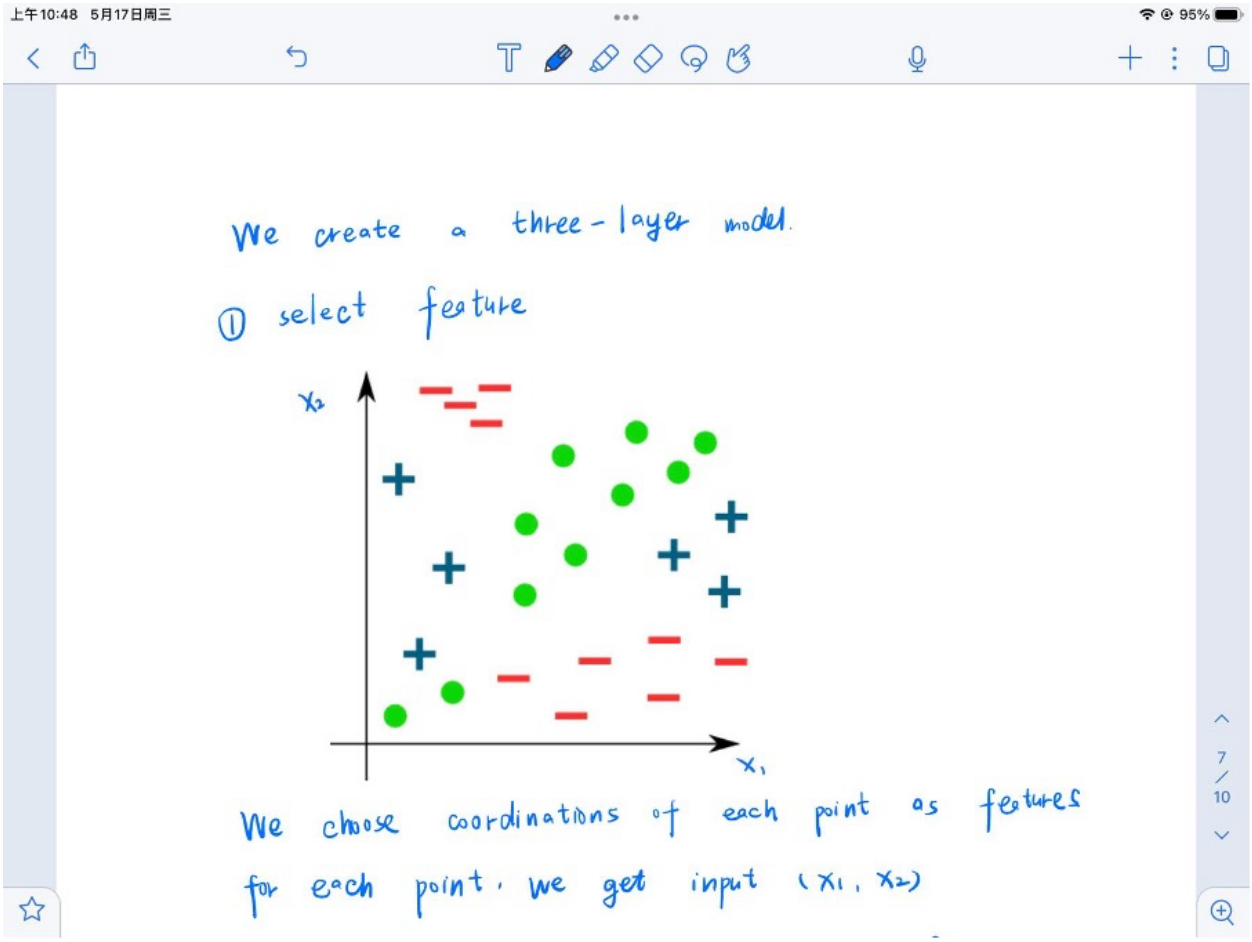
Exercise 2

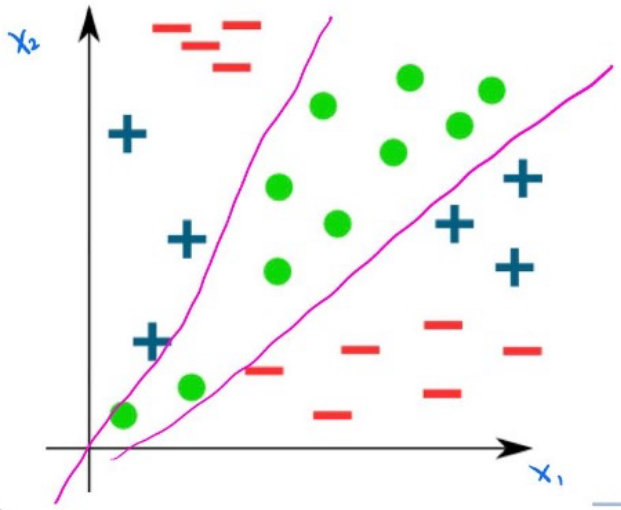
Group: Sevde Yanik & Wen Jiang & Ian D. Fichtner

1 Hand-Crafted Network



① the first lyer, we use logical OR funtion

label = { 1 if it's green
0 else



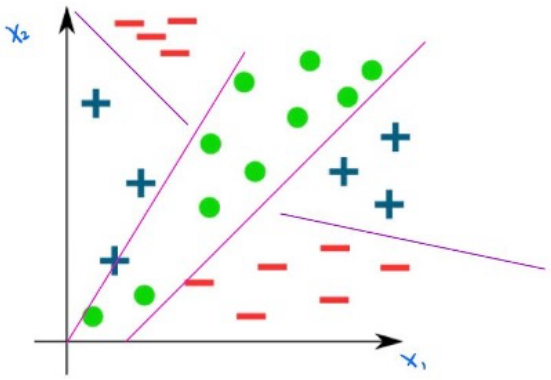
so we get two lines to distinguish green points with others

③ the second layer we use masked logical OR.

$$label2 = \begin{cases} 1 & \text{if it's red} \\ 0 & \text{else} \end{cases}$$

label 1 is regarded as C in function $g(z:c)$

so we can get classification:



now we get a zero training error classification

④ the last layer we use "perfect match" to get the final output.

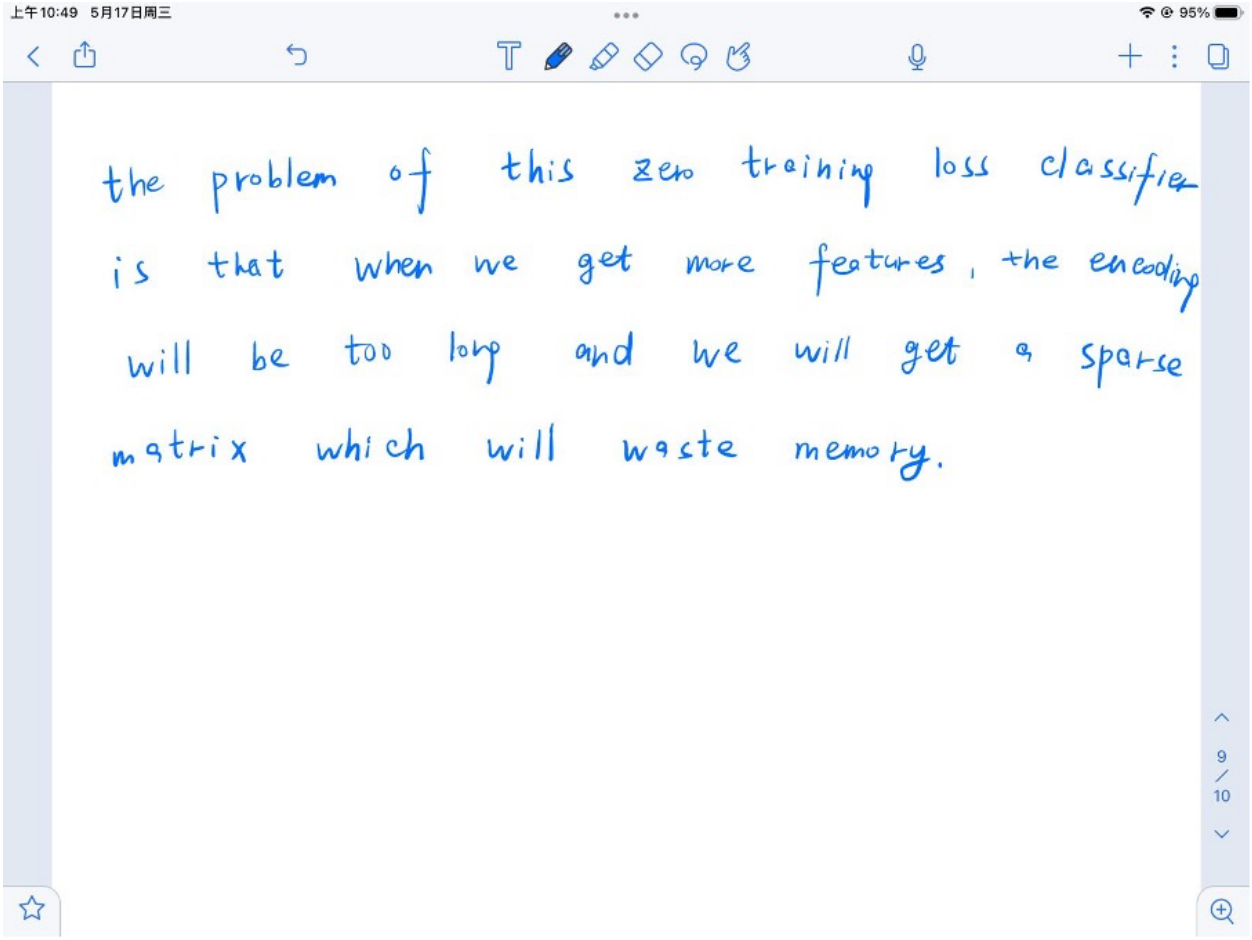
	red minus	blue plus	green circle
$f(z)$	0	0	1
$g(z)$	1	0	0
$h(z)$	0	1	0

So finally we get a one-hot encoding:

class 1: (0, 1, 0)

class 2: (0, 0, 1)

class 3: (1, 0, 0)



2 Linear Activation Function

Since we assume φl is identity function, $Z_1 = \tilde{Z}_1$

$Z_0 = X$

$\tilde{Z}_1 = Z_0 \cdot B_1 + b_1$

$= X \cdot B_1 + b_1$

$\tilde{Z}_2 = Z_1 \cdot B_2 + b_2$

$= (X \cdot B_1 + b_1) \cdot B_2 + b_2$

$= X \cdot B_1 \cdot B_2 + b_1 \cdot B_2 + b_2$

Let's say for the L-layer network B' is the product of matrices the $B' = B_1 \cdot B_2 \cdot \ldots \cdot B_L$ and b_L is the sum of biases.

Then, $Z_L = X \cdot B' + b_L$

Which essentially means the output of the L-layer network is equivalent to a 1-layer network.