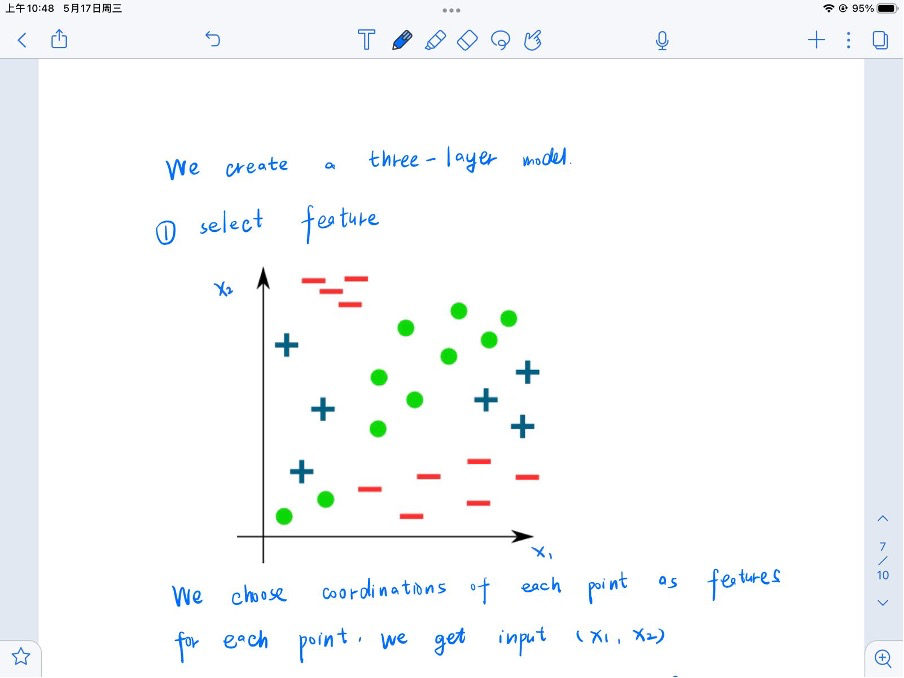
**Exercise 2**

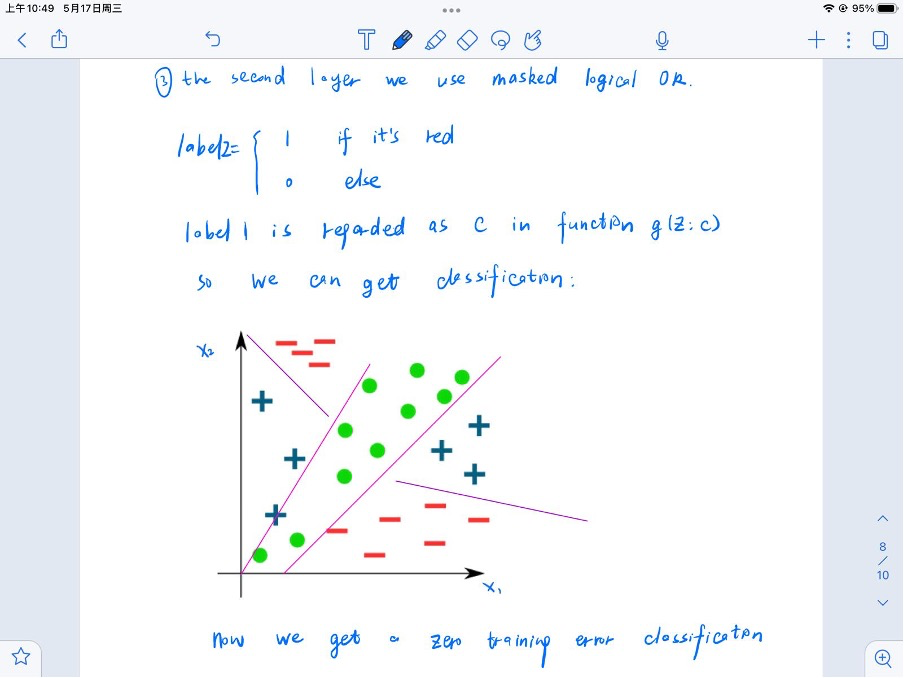
*Group: Sevde Yanik & Wen Jiang & Ian D. Fichtner*

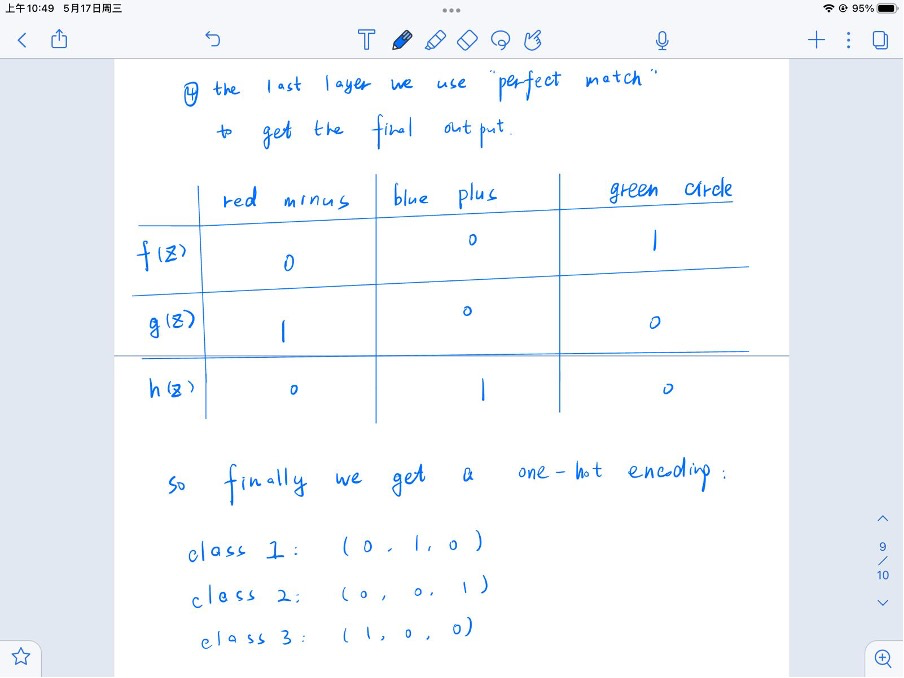
**1**      **Hand-Crafted Network**

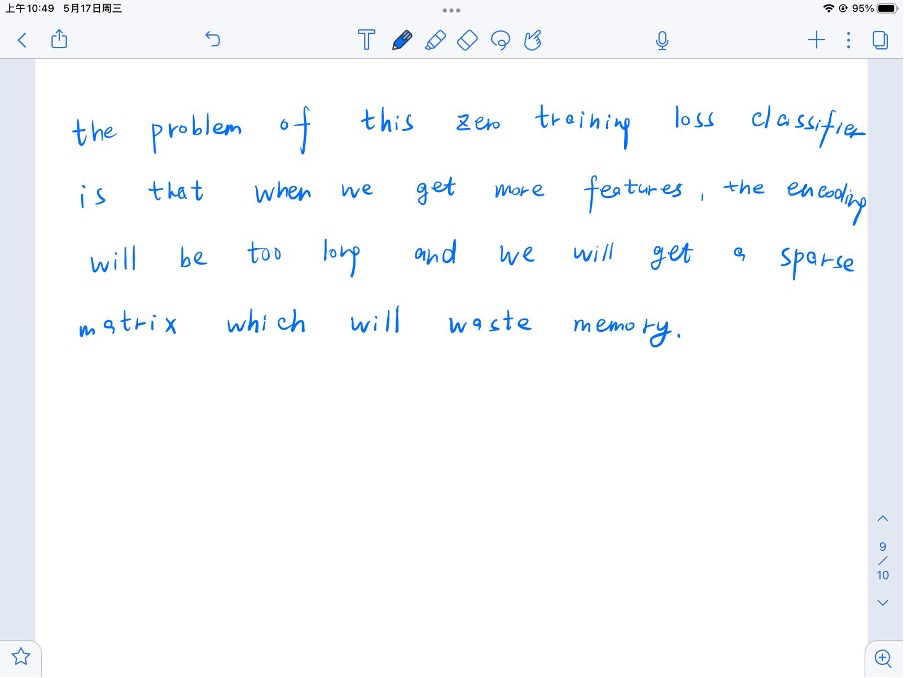
****

**A screenshot of a whiteboard with text

Description automatically generated with low confidence**

****

****

****

**2**      **Linear Activation Function**

Since we assume φl is identity function, Z l = Z ̃l

Z0 = X

Z ̃1  =  Z0 · B1 + b1

          =  X · B1 + b1

Z ̃2  =  Z1 · B2 + b2

         = (X · B1 + b1)  · B2 + b2

      = X · B1 · B2  +  b1 · B2  + b2

Let’s say for the L-layer network B’ is the product of matrices the B’ = B1 · B2 · … · BL     and bL is the sum of biases.

Then, ZL = X · B’ + bL

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