Quiz 2
Week 3, Sep./21/2022
CS 280: Fall 2022
Instructor: Lan Xu

Name: _____
On your left: _____
On your right: _____

Instructions:
Please answer the questions below. Show all your work. This is an open-book test. NO discussion/collaboration is allowed.

Problem 1. (10 points)
Consider the perceptron algorithm and let us re-write the weight updates as follows:

$$\begin{aligned}
\text{Initialization}: w_1^+ &= w_1^- = 0 \\
\text{Mistakes on positive}: w_{t+1}^+ &= w_t^+ + x \\
\text{Mistakes on negative}: w_{t+1}^- &= w_t^- - x \\
\text{Weight update}: w_{t+1} &= w_{t+1}^+ + w_{t+1}^-
\end{aligned}$$

If the inputs are images from two categories: apple (positive) and banana (negative). What would the final weights of w_T^+ and w_T^- look like as an image after T iterations, and why? What about w_T?

$T=1$, $w_T = w_1^+ + w_1^- = 0$ and it makes mistake.

$T>1$, if an apple image x_{apple} is given, $w_{t+1}^+ = w_t + x_{apple}$, $w_{t+1} = w_t + x_{apple}$ (+4)

if an banana image x_{banana} is given, $w_{t+1}^- = w_t - x_{banana}$, $w_{t+1} = w_t - x_{banana}$

Thus, w_T^+ will looks like an apple after iterations. (+3)

Similarly, w_T^- will look like an inverse banana.

Therefore, w_T will look like an image of an apple substract a banana. (+3)

说明: ① 有迭代过程但没答出apple / banana的可酌情给分.

② 只有答案没有过程的要扣过程分.

Problem 2. (10 points)

Consider a multiclass logistic regression with L1 regularization as follows:

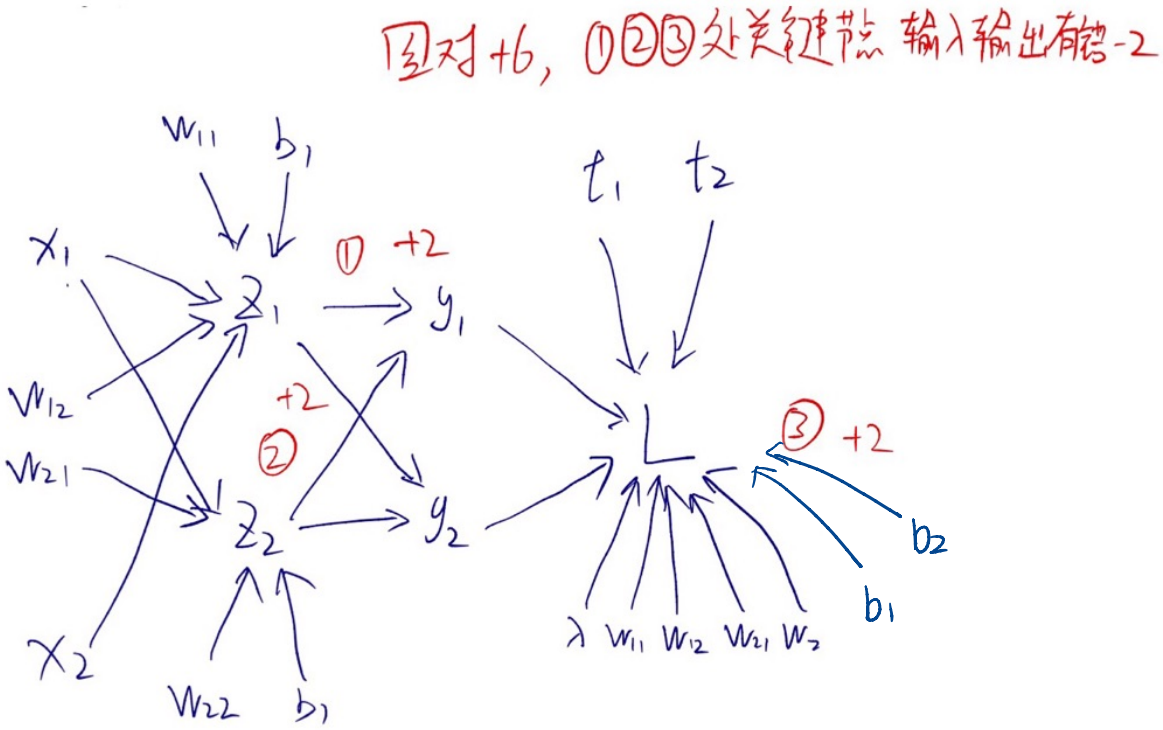
$$z_l = \sum_{j=1}^{2} w_{lj} x_j + b_l, l = 1, 2$$

$$y_k = \frac{e^{z_k}}{\sum_l e^{z_l}}, k = 1, 2$$

$$\mathcal{L} = -\sum_k t_k \log y_k + \lambda \sum_{i=1}^{2} \sum_{j=1}^{2} ||w_{ij}||_1 + \gamma \sum_{l=1}^{2} ||b_l||_1$$

Draw a computational graph for this network and its loss. Note each node should be a scalar in this graph.

Write down the forward pass based on the graph you have built.



$z_1 = w_{11} x_1 + w_{12} x_2 + b_1$

$z_2 = w_{21} x_1 + w_{22} x_2 + b_2$

$y_1 = \frac{e^{z_1}}{\sum_{i=1}^{2} e^{z_i}}$ $y_2 = \frac{e^{z_2}}{\sum_{i=1}^{2} e^{z_i}}$

L = 照抄题干即可 ...

+4 forward 过程

2