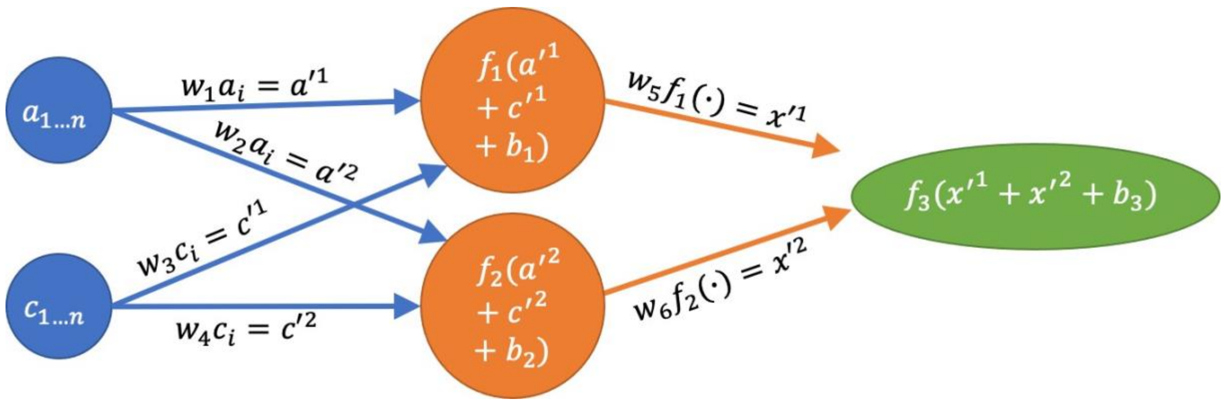
CSE514, Fall 2021, HW 3 Name: Anh Le Student ID: 488493

Note: This homework is worth a total of 15 points

Q1: Given this ANN structure:



And the following parameter/function definitions:

W = [-15, -3, -2, 4, 1, 10]   B = [4, 1, -0.5]

$f_1(x) = f_2(x) = \max(0.1x, x)$ $f_3(x) = x^2$

What are the intermediate and/or output values for the following data points?

Q1a (3pt): Data point: a = 0.5, c = 0.5

a'^1 value: -7.5

a'^2 value: -1.5

c'^1 value: -1

Q1b(3pt): Data point: a = 1, c = 0

c'^2 value: 0

f_1 value: max (0.1(-15 + 0 + 4), (-15 + 0 + 4)) = –1.1

f_2 value: max (0.1(–3 + 0 + 1), (–3 + 0 + 1)) = –0.2

Q1c (3pt): Data point: a = 0, c = 1

x'^1 value: 1 * max (0.1(0 – 2 + 4), (0 – 2 + 4)) = 2

x'^2 value: 10 * max (0.1(0 + 4 + 1), (0 + 4 + 1)) = 50

f_3 value: $(2 + 50 - 0.5)^2 = 2652.25$

Q2: Use the same ANN structure as Q1.

Write out the formulas for calculating the partial derivatives of the loss function with respect to the parameters asked for below.

Use SSE/2 as the loss function.

Example: parameter w_5

$$\frac{\partial L}{\partial w_5} = \frac{\partial L}{\partial f_3} * \frac{\partial f_3}{\partial x'^1} * \frac{\partial x'^1}{\partial w_5}$$

$$= \frac{\partial}{\partial f_3}\left(\frac{1}{2}\sum_i^n (f_3 - observed)^2 * \frac{\partial}{\partial x'^1}(x'^1 + x'^2 + b_3)^2 * \frac{\partial}{\partial w_5}(w_5 f_1)\right)$$

$$= \sum_i^n (f_3 - observed) * 2(x'^1 + x'^2 + b_3) \; * \; f_1$$

Q2a (2pt): parameter b_3

$$\frac{\partial L}{\partial b_3} = \frac{\partial L}{\partial f_3} \times \frac{\partial f_3}{\partial b_3} = \frac{1}{2}\sum (f_3 - observed)^2 \times \frac{\partial f_3}{\partial b_3}$$

$$= \sum (f_3 - observed) \times \frac{\partial}{\partial b_3}(x'^1 + x'^2 + b_3)^2 = \sum (f_3 - observed)\, 2(x'^1 + x'^2 + b_3)$$

Q2b (2pt): parameter w_6

$$\frac{\partial L}{\partial w_6} = \frac{\partial L}{\partial f_3} \times \frac{\partial f_3}{\partial x'^2} \times \frac{\partial x'^2}{\partial w_6}$$

$$= \frac{\partial}{\partial_3}\left(\frac{1}{2}\sum (f_3 - observed)^2 \times \frac{\partial}{\partial x'^2}(x'^1 + x'^2 + b_3)^2 \times \frac{\partial}{\partial w_6}(w_6 f_2)\right)$$

$$= \sum (f_3 - observed) \times 2(x'^1 + x'^2 + b_3)\, f_2$$

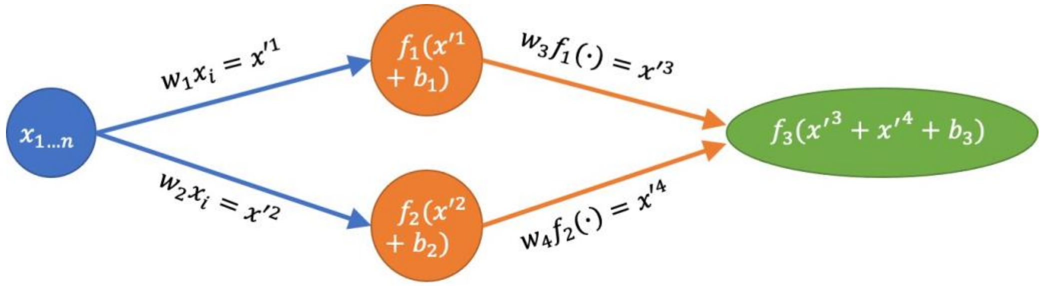
Q3: Given an image of 22x31 pixels, and a filter of size 3x3, what would be the size of your feature map given:

Q3a (1pt): No padding, stride = 1

$$\frac{22 - 3 + 2 \times 0}{1} + 1 = 20 \qquad \frac{31 - 3 + 2 \times 0}{1} + 1 = 29 \Rightarrow 20 \times 29$$

Q3b (1pt): Padding of 1 on all sides of the image, stride = 2

$$\frac{22 - 3 + 2 \times 1}{2} + 1 = 11 \qquad \frac{31 - 3 + 2 \times 1}{2} + 1 = 16 \Rightarrow 11 \times 16$$



Loss function: SSE/2 Data = ((.5, .5), (1, 0), (0, 1))

$W = [-25, -3, -2, 4]$ $f_1(x) = f_2(x) = \max(0.1x, x)$

$B = [4, 1, -0.5]$ $f_3(x) = x$

Extra credit:
(1 bonus pt.)
Calculate the new W and B values using backpropagation, and a learning rate = 0.01

$$\frac{\partial L}{\partial b_3} = \frac{\partial L}{\partial f_3} \times \frac{\partial f_3}{\partial b_3} = \sum (f_3 - \text{observed}) \times \underbrace{\frac{\partial}{\partial b_3}(x'^3 + x'^4 + b_3)}_{=1}$$

$$= \left[(1 - 0.5) + (2.9 - 0) + (-4.5 - 1) \right] = -2.1$$

$$\rightarrow b_3' = b_3 - \alpha \frac{\partial L}{\partial b_3} = -0.5 - 0.01 \times -2.1 = -0.479$$

$$\frac{\partial L}{\partial w_3} = \frac{\partial L}{\partial f_3} \times \frac{\partial f_3}{\partial x'^3} \times \frac{\partial x'^3}{\partial w_3} = \sum (f_3 - \text{observed}) \underbrace{\left[\frac{\partial}{\partial x'^3}(x'^3 + x'^4 + b_3)\right]}_{=1} \underbrace{\left[\frac{\partial}{\partial w_3}(w_3 f_1)\right]}_{=f_1}$$

$$= \sum (f_3 - \text{observed}) f_1 = \left[(1-0.5) \cdot 0.85 + (2.9 - 0) \cdot 2.1 + (-4.5 - 1) 4 \right] = -28.515$$

$$\rightarrow w_3 = w_3 - \alpha \frac{\partial L}{\partial w_3} = -2 - 0.01 \times (-28.515) = -1.71485$$

$$\frac{\partial L}{\partial w_4} = \frac{\partial L}{\partial f_3} \times \frac{\partial f_3}{\partial x'^4} \times \frac{\partial x'^4}{\partial w_4} = \sum (f_3 - \text{observed}) \underbrace{\left[\frac{\partial}{\partial x'^4}(x'^3 + x'^4 + b_3)\right]}_{=1} \underbrace{\left[\frac{\partial}{\partial w_4}(w_4 f_2)\right]}_{=f_2}$$

$$= \sum (f_3 - \text{observed}) f_2 = \left[(1-0.5) \cdot 0.05 + (2.9 - 0) \cdot 0.2 + (-4.5 - 1) 1 \right] = -6.105$$

$$\rightarrow w_4 = w_4 - \alpha \frac{\partial L}{\partial w_4} = 4 - 0.01(-6.105) = 4.06105$$

$$\frac{\partial L}{\partial b_1} = \frac{\partial L}{\partial f_3} \times \frac{\partial f_3}{\partial x'^3} \times \frac{\partial x'^3}{\partial f_1} \times \frac{\partial f_1}{\partial b_1} = \sum (f_3\text{-observed})\, 1\, \overbrace{\frac{\partial}{\partial f_1}(w_3 f_1)}^{= w_3} \frac{\partial f_1}{\partial b_1}$$

$$= \left[(1-0.5)\text{-}2(0.1) + (2.9-0)\text{-}2(0.1) + (-4.5-1)\text{-}2(1)\right] = 10.32$$

$$\rightarrow b_1 = b_1 - \alpha \frac{\partial L}{\partial b_1} = 4 - 0.01 \times 10.32 = 3.8968$$

$$\frac{\partial L}{\partial b_2} = \frac{\partial L}{\partial f_3} \times \frac{\partial f_3}{\partial x'^4} \times \frac{\partial x'^4}{\partial f_2} \times \frac{\partial f_2}{\partial b_2} = \sum (f_3\text{-observed})\, 1\, \overbrace{\frac{\partial}{\partial f_2}(w_4 f_2)}^{= w_4} \frac{\partial f_2}{\partial b_2}$$

$$= \left[(1-0.5)4(0.1) + (2.9-0)4(0.1) + (-4.5-1)4(1)\right] = -20.64$$

$$\rightarrow b_2 = b_2 - \alpha \frac{\partial L}{\partial b_2} = 1 - 0.01(-20.64) = 1.2064$$

$$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial f_3} \times \frac{\partial f_3}{\partial x'^3} \times \frac{\partial x'^3}{\partial f_1} \times \frac{\partial f_1}{\partial x'^1} \times \frac{\partial x'^1}{\partial w_1}$$

$$= \sum (f_3\text{-observed})\, 1\, (w_3) \frac{\partial f_1}{\partial x'^1} \times \underbrace{\frac{\partial}{\partial w_1}(w_1 x_i)}_{x_i}$$

$$= \left[(1-0.5)\text{-}2(0.1)0.5 + (2.9-0)\text{-}2(0.1)1 + (-4.5-1)\text{-}2(1)0\right] = -0.63$$

$$\rightarrow w_1 = w_1 - \alpha \frac{\partial L}{\partial w_1} = -25 - 0.01(-0.63) = -24.9937$$

$$\frac{\partial L}{\partial w_2} = \frac{\partial L}{\partial f_3} \times \frac{\partial f_3}{\partial x'^4} \times \frac{\partial x'^4}{\partial f_2} \times \frac{\partial f_2}{\partial x'^2} \times \frac{\partial x'^2}{\partial w_2}$$

$$= \sum (f_3\text{-observed})\, 1\, (w_4) \frac{\partial f_2}{\partial x'^2} \times \underbrace{\frac{\partial}{\partial w_2}(w_2 x_i)}_{= x_i}$$

$$= \left[(1-0.5)4(0.1)0.5 + (2.9-0)4(0.1)1 + (-4.5-1)4(1)0\right] = 1.26$$

$$\rightarrow w_2 = w_2 - \alpha \frac{\partial L}{\partial w_2} = -3 - 0.01 \times 1.26 = -3.0126$$

$$\rightarrow \begin{cases} \text{new W values: } [-24.9937, -3.0126, -1.71485, 4.06015] \\ \text{new B values: } [3.8968, 1.2064, -0.479] \end{cases}$$

(1 bonus pt.)
Calculate the error with the current W and B values, and then the error with the new W and B values.

Current error $= \frac{1}{2}\sum(f_3 - \text{observed})^2 = \frac{1}{2}\left[(1-0.5)^2 + (2.9-0)^2 + (-4.5-1)^2\right] = \boxed{19.455}$

0.5

$w_1x_1 = -12.49685$

$\max[0.1(-12.49685 + 3.8968), (-12.49685 + 3.8968)]$

$\hookrightarrow f_1 = -0.86$

$w_2x_1 = -1.5063$

$\max[0.1(-1.5063 + 1.2064), (-1.5063 + 1.2064)]$

$\hookrightarrow f_2 = -0.03$

$w_3f_1 = 1.474771$

$w_4f_2 = -0.1218045$

$f_3 = 1.474771 - 0.1218045 - 0.479 = 0.874$

1

$x_1' = -24.9937$ → $f_1 = -2.10969$ → $x'^3 = 3.6178$

$x_2' = -3.0126$ → $f_2 = -0.18062$ → $x'^4 = -0.73334$

$f_3 = 2.40546$

0

$x_1' = 0$ → $f_1 = 3.8968$ → $x'^3 = -6.6824$

$x_2' = 0$ → $f_2 = 1.2064$ → $x'^4 = 4.8982$

$f_3 = -2.2633$

New Error $= \frac{1}{2}\left[(0.874-0.5)^2 + (2.40546-0)^2 + (-2.2633-1)^2\right] = \boxed{8.2876}$