Homework 8

Instructions
This homework contains **1** concept and **4** programming questions. In MS word or a similar text editor, write down the problem number and your answer for each problem. Combine all answers for concept questions in a single PDF file. Export/print the Jupyter notebook as a PDF file including the code you implemented and the outputs of the program. Make sure all plots and outputs are visible in the PDF.

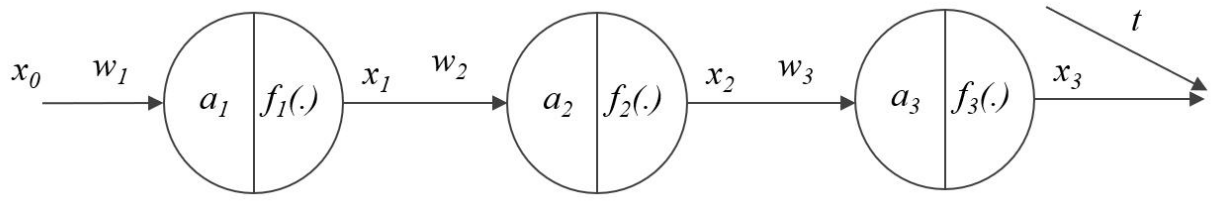
Combine all answers into a single PDF named andrewID_hw8.pdf and submit it to Gradescope before the due date. Refer to the syllabus for late homework policy. Please assign each question a page by using the "Assign Questions and Pages" feature in Gradescope.

Here is a breakdown of the points for programming questions:

Question	**Points**
Concept 1	12
M8-L1-P1	12
M8-L2-P1	12
M8-L2-P2	12
M8-HW1	72
Total	**120**
Bonus	6

Problem 1

Consider the following network, with $x_0 = 2$, $w_1 = -1$, $w_2 = 3$, $w_3 = 7$, and linear (identity) activation functions.
Compute $\partial L/\partial w_3$, $\partial L/\partial w_2$, $\partial L/\partial w_1$ provided that t = -40
$L = \frac{1}{2} e^T e$



$a_1 = w_1 \cdot x_0 = -1 \cdot 2 = -2$

$x_1 = f_1(a_1) = -2$

$a_2 = w_2 \cdot x_1 = 3 \cdot -2 = -6$

$x_2 = f_2(a_2) = -6$

$a_3 = w_3 \cdot x_2 = 7 \cdot -6 = -42$

$x_3 = f_3(a_3) = -42$

$e = x_3 - t = -2 \quad \Rightarrow \quad L = \frac{1}{2} \cdot (-2) \cdot (-2) = 2$

$$\frac{\partial L}{\partial w_3} = \frac{\partial L}{\partial x_3} \frac{\partial x_3}{\partial a_3} \frac{\partial a_3}{\partial w_3}$$

$= -2 \cdot 1 \cdot -6 = 12$

$$\frac{\partial L}{\partial w_2} = \frac{\partial L}{\partial x_3} \frac{\partial x_3}{\partial a_3} \frac{\partial a_3}{\partial x_2} \cdot \frac{\partial x_2}{\partial a_2} \cdot \frac{\partial a_2}{\partial w_2}$$

$= -2 \cdot 1 \cdot 7 \cdot 1 \cdot (-2)$

$= 28$

$$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial x_3} \frac{\partial x_3}{\partial a_3} \cdot \frac{\partial a_3}{\partial x_2} \cdot \frac{\partial x_2}{\partial a_2} \cdot \frac{\partial a_2}{\partial x_1} \cdot \frac{\partial x_1}{\partial a_1} \cdot \frac{\partial a_1}{\partial w_1}$$

$= -2 \cdot 1 \cdot 7 \cdot 1 \cdot 3 \cdot 1 \cdot 2$

$= -84$