Name: Kaley Nguyen Email: kaleynn2@illinois.edu

Question 1. What are the main messages you learned from this chapter?

A linear transformation of the input can become a non-linear transformation with activation function. Leaky ReLU appears to be the greatest option so far for the activation function because it overcomes both the problems of sigmoid/gradient tanh's vanishing gradient issue and ReLU's zero values when x is negative.

Question 2. What are the typos?

- Page 41: and RNN for clinical event prediction from **patients'** longitudinal EHR.
- Page 45: However, despite its simplicity, ReLU is still a nonlinear function**,** which is essential for **the** neural network learning.
- Page 45: the regular ReLU function may be stuck at zero when x **is** negative, limiting the learning capacity of the neuron.
- Page 46: Compared to L2 loss, the L1 loss is more robust in **a sense** that it is more resistant to outliers in the data.

Question 3. Which part do you want to improve in this chapter?

In this section, I'd want to include the L-norm and why the L1 loss is more resistant to outliers than L2.

Question 4. Which types of health data do you think can be benefited most by DL methods? Why?

The majority of the unstructured data has been processed using DL techniques.

CNN: can be used to automatically classify diseases and medical images. RNN: can be utilized to simulate clinical event sequences, such as intervention/treatment results across time, as well as sequential data, such as text and time series data.

Question 5. Which types of health data do you think can be benefited least by DL methods? Why?

Since it takes too long for a prediction to take effect, DL techniques cannot handle real-time EKG/EEG prediction.