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

- Graphs are standard representation of pairwise relations in a lot of healthcare applications. Some graph related concepts are introduced such as:

  > ***Normalization****: $$P = D^{-1}A$$ instead of using the adjacency matrix A.*
  > ***Spectral Analysis****: The eigen-decomposition on a Laplacian matrix L is referred to as spectral analysis. In other words, $$L = Q\wedge Q^T$$ where $$\wedge$$ is a diagnonal matrix with eigenvalues and Q are the eigenvector matrix.*

  Several common graph tasks:
- Node classification: classify each node in a graph into a set of targeted labels.
- Link prediction: given a graph of unknown drug-to-drug protein interations, we can estimate the unknown drug-to-drug protein interactions.
- Community detection: can help to locate the subgraph / sub-disease clusters within the graph.
- Graph property prediction: given an existing molecular graph and the associated drug properties, we can estimate those drug properties on the other unknown molecules.
- Graph generation: also known as molecule generation or the process of creating new graphs based on the exisitng graphs.

Question 2. What related resources (book, paper, blog, link) do you recommend your classmates to checkout?

The original paper of decagon is one of the few papers that I would like to understand 3 years ago. I am really happly to learn about the same paper in the class.

http://snap.stanford.edu/decagon/

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

I'd like to see more images of the networks, like a bipartite graph or a small capture of SNOMED.

Question 4. What is the main difference between graph convolutional network and graph attention network?

- graph convolutional network: GCN is an approximation of the original GNN. In GCN, a node is represented as a function of its neighbors; similarly, adjacent pixels in an image are localize. GCN gathers data from a node's immediate area iteratively to create improved embeddings for the node. The nodes share the same W weight matrix.

- graph attention network: GAT uses the attention mechanism on graphs, or applying a dynamic weight matrix W. GAT allows different weights to different nodes in the neighborhood.