Name: Kaley Nguyen Email: kaleynn2@illinois.edu

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

Using embedding techniques like Word2Vec, Glove, t-SNE, Med2Vec, and MIME, all learning concept representations seek to clarify the logitudinal EHR and produce some useful clinical applications. All of the embedding methods have greatly improved both performance and prediction time because the one-hot encoding technique is no longer effective and computationally expensive.

Question 2. Which part do you want to improve in this chapter?

Glove is mentioned in the beginning of the chapter but it does not have any associated sub-chapter.

Question 3. What related topics do you suggest adding in this chapter?

As an additional autonomous learning technique that can support t-SNE, I would like to mention UMAP.

Question 4. What is the key difference between Med2Vec and Word2Vec?

- Word2Vec: Word2Vec attempts to learn a vector representation, also known as embedded vectors, for each word so that some closely related diagnoses can be closed to each other rather than dissimilar diagnoses. One-hot encoding vocabulary construction of N unique diagnoses using N-dimensional vectors has some limitations of similar pairwise distance. Word2Vec outperforms one-hot encoding in terms of having a higher AUC while having a significantly smaller vocabulary construction, assuming that each word's meaning is decided by its neighbors.
- Med2Vec: Med2Vec seeks to distinguish the hierarchical differences between the sequential visits and co-occurence of medical codes because medical codes do not retain the sequential pattern of visit progression. Because both Word2Vec and Med2Vec attempt to predict the easrly onset of medical conditions, they are comparable. Given that it takes into consideration the sequential pattern in the visits, Med2Vec scores better than Word2Vec in the paper mentioned in the textbook.

Question 5. What is the key difference between Med2Vec and MiME?

- Med2Vec: Med2Vec uses the two-level structure of longitudinal EHR, which are visits and medical codes.
- MiME: MiME employs the longitudinal EHR's three-level framework, which consists of codes for medical diagnosis, visits, and treatment (including procedures and medications). For instance, there is a connection between the diagnosis and treatment codes for a particular diagnosis. To get the embedded vectors, all of the diagnosis objects from a single visit are used. In every dataset, MiME scores better than other baseline models, with the smaller datasets showing the largest differences from the baseline models.

Typos:

Page 66: By maximizing the softmax score of the inner product of embeddings between each target word and its neighboring words, ~~By maximizing this softmax function,~~ Word2Vec learns real-valued vectors that capture the co-occurance relations between words.