Saaif Ahmed

saaifza2@illinois.edu

Chapter 5 Reflection

1. Please give a brief summary of the chapter?

The chapter outlines the attention mechanism's role in sequence-to-sequence (seq2seq) prediction tasks, such as generating future medication sequences based on a patient's historical disease diagnoses. Recent advancements in neural network models have significantly enhanced translation systems such as Google Translate. Similarly, in healthcare, these models play a crucial role in transforming various clinical information types, enabling tasks like predicting medication prescriptions based on historical diagnoses and forecasting future health conditions using time-series data like vital signs and lab results.

A simple RNN-based approach is introduced before discussing the attention mechanism. Traditional RNN models capture sequential information but require separate models for encoding the source sequence and generating the dynamic target sequence, as demonstrated in this study. The traditional approach of using two Recurrent Neural Networks for processing source sentences in seq2seq models faces limitations. To overcome these challenges, attention-based seq2seq models are developed, which generate tokens one at a time while considering both previously decoded words and corresponding tokens in the source sequence, like human processing.

The chapter then gives us information on the different attention functions and then goes into a practical example.

2. What improvements do you want to see in this chapter? Please elaborate on them

With the sheer amount of code in this section I would highly suggest putting the code side by side and using the whole page. It is very difficult to read through the gaps in the pages. And follow along with the code. I would also like the math sections to be a bit larger in font size so that it is clear what is going on.

3. What are the typos in this chapter?

- Page 77 "sequence o f diagnoses" has an extra space in the "of"

Otherwise no other typos were noticed.

4. Which part of the chapter do you like most

I quite liked the explanation portions of this chapter, as it gave very good detail into how the seq2seq models can be utilized. The math portions were also good despite me not being able to understand them immediately.

5. What are the most useful things you learned from this chapter?

The most useful part of this chapter was the code blocks and the multiple examples given to the readers to show how this works. The example at the end of the chapter which walks through an entire example is perfect for explaining how one uses PyTorch to make an RNN with an attention mechanism. And the makings of that were explained well all throughout the chapter.

6. Could you find at least one research papers that use attention models for handling healthcare predictive tasks? Use one sentence to summarize the paper and add citation.

The study by Kaji et. Al [1] trained an LSTM RNN and incorporated an attention mechanism to aid in the prediction of daily sepsis and other issues in patients that were in an Intensive Care Unit over the course of two weeks.

References

[1] Kaji DA, Zech JR, Kim JS, Cho SK, Dangayach NS, Costa AB, et al. (2019) An attention
based deep learning model of clinical events in the intensive care unit. PLoS ONE 14(2):
e0211057. https://doi.org/10.1371/journal.pone.0211057

[2] Xiao, C., Sun, J. (2021). Introduction. In: Introduction to Deep Learning for Healthcare.
Springer, Cham. https://doi-org.proxy2.library.illinois.edu/10.1007/978-3-030-82184-5_1