

Saaif Ahmed

saaifza2@illinois.edu

Chapter 6 Reflection

1. Please give a brief summary of the chapter?

The attention mechanism is a crucial concept in neural network methods, utilized for enhancing prediction accuracy and model interpretability. Transformers serve as the basis for various advanced neural network applications, including large language models. Transformer models find use in tasks such as rare disease detection, clinical outcome prediction, and synthetic patient data generation.

The chapter then delves into the attention mechanism's origin within the sequence-to-sequence (Seq2Seq) prediction setting. Traditional RNN-based models faced challenges in training for Seq2Seq problems due to intensity and optimization problems. The attention mechanism, introduced separately of sequence dependencies from RNNs, overcame this. Transformers have become foundational in models like BERT and large language models such as ChatGPT, demonstrating versatility beyond text and potential for processing multi-modal data.

Towards the chapter's end, two practical examples of using transformers for clinical prediction tasks are provided.

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

While it presents a significant challenge, exploring improved visual representations of convolution could be good. The current graphs and images are easy to implement and decent for individuals who can mentally visualize mathematical concepts effectively. However, given the graphical nature of the topic, investing in more compelling graphics would be highly valuable. A set of meticulously crafted high-quality graphics illustrating the process of the RNN would greatly enhance the clarity of the lesson. This is particularly important for individuals who struggle with matrices of numbers, onto other matrices.

3. What are the typos in this chapter?

- Pg 127 “Specifically, let us delve into the drug recommendationmimic3f n task function.” has typos

Otherwise no other typos were noticed.

4. Which part of the chapter do you like most

I quite liked the efficiency improvements section showing how these developments drastically improve over the old design.

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

The code blocks and numerous examples provided in this chapter were particularly beneficial, offering readers practical demonstrations of the concepts discussed. The final example, which guides through a complete process, is especially effective in illustrating how PyTorch can be utilized to implement an RNN with an attention mechanism. The chapter effectively builds up to this demonstration by thoroughly explaining the components and workings involved.

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

ExBEHRT[1] Extends upon the research paper shown in the example. It expands upon the inputs of BEHRT.

References

- [1] Maurice Rupp, Oriane Peter, and Thirupathi Pattipaka. 2023. ExBEHRT: Extended Transformer for Electronic Health Records. In Trustworthy Machine Learning for Healthcare: First International Workshop, TML4H 2023, Virtual Event, May 4, 2023, Proceedings. Springer-Verlag, Berlin, Heidelberg, 73–84. https://doi.org/10.1007/978-3-031-39539-0_7
- [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