**Graded Discussion 5**

Transformer models have had a significant impact on improving the accuracy and flexibility of the NLP models. In fact, in some cases, transformer models with a large number of learnable parameters have achieved or exceeded human capabilities at some tasks.

Consider an NLP application where a transformer model can be used. What is the application? What data of sufficient volume could you hypothetical use to train your model?  What difficulties do you anticipate with training this model for the application?

Two main applications come to mind when I consider where a transformer model might be use in an NLP setting are: Book Translation & Question Answering problems related to highly technical domains.

As for the first one the situation where I envision a Transformer model helping is in quicker less expensive turnaround times for books to be translated into multiple languages. This from an anthropology and sociology perspective could help in the cross pollination of knowledge. Right now, this effort is based on revenue as the investment is quite high and just high-profile books (booksellers) get to be translated into specific languages putting a barrier for knowledge into many domains. As an example, last year, I read “Theology of the Pain of God” by Kazoh Kitamori, which now days is a classic book in non-wester Christian theology with original. This book took 20 years to be translated into English, which might not be a long time compared to other books, but when it was finally translated to English brough with it a revolution into the west theologian’s community as a new perspective was exposed.

The other case that I mentioned is the Question – Answering problem for Highly Technical domains, which is in a way a translation problem from a problem language domain to a solution language domain. This to solve the problem of areas where similar problems are addressed by different groups but with little to no communication between these groups (as could happen in big companies such as Boeing or at greater scale the scientific community).

In both cases the data in turn that could be used would be the available books on the domain that we are trying to translate. For example, if the book is in finance or self-improvement then look for contemporary books on the field to capture the language used this for the encoder and then for the decoder do the same but using books in the target language.

As for the Question – Answering application, gathering as many examples as possible from written problems and their respective solutions for the entire company historical data would be the encoder and decoder parts respectively.

In both cases the main issue would come from the amount of data available to train the model as Transformer Models require a huge amount of data, specially taking into account that we need to have relevant examples for translation to have a successful encoding-decoding model. On the bright side we could potentially use some pre-trained models such as GPT-3 (Generative Pre-Trained Transformer Model – Autoregressive Language Model) and just add extra layers of training with the data available.

References:

“The Transformer Attention Mechanism” <https://machinelearningmastery.com/the-transformer-attention-mechanism/>

“All you need to know about Attention and Transformers” <https://towardsdatascience.com/all-you-need-to-know-about-attention-and-transformers-in-depth-understanding-part-1-552f0b41d021>

“Attention and Transformers Modes” <https://towardsdatascience.com/attention-and-transformer-models-fe667f958378>

“4 Reasons Transformers Models are Optimal for NLP <https://www.eweek.com/big-data-and-analytics/reasons-transformer-models-are-optimal-for-handling-nlp-problems/>

“Will Transformers Take Over AI” <https://www.quantamagazine.org/will-transformers-take-over-artificial-intelligence-20220310/>

“Machines Beat Humans on a Reading Test. But Do They Understand?” <https://www.quantamagazine.org/machines-beat-humans-on-a-reading-test-but-do-they-understand-20191017/>

4 Reasons Transformers Models are Optimal for NLP:

Since their initial development in the seminal AI research paper [Attention Is All You Need](https://arxiv.org/abs/1706.03762), transformer-based architectures have completely redefined the field of [Natural Language Processing (NLP)](https://www.webopedia.com/definitions/nlp/) and set the state of the art for numerous AI benchmarks and tasks.

Will Transformers Take Over AI:

Even if transformers can help unite and improve the tools of AI, emerging technologies often come at a steep cost, and this one is no different. A transformer requires a higher outlay of computational power in the pre-training phase before it can beat the accuracy of its conventional competitors.

It also suggests the tantalizing prospect of some hybrid architecture that draws on the strengths of transformers in ways that today’s researchers can’t predict. “Perhaps we shouldn’t rush to the conclusion that the transformer will be the final model,” Wang said. But it’s increasingly likely that the transformer will at least be a part of whatever new super-tool comes to an AI shop near you.