SP0061 Reflection Assignment

# Content

I found the content in this module extremely brief and surface-level, which is insufficient for a good understanding of a topic. The technical aspects of a topic are usually covered within a week, and usually so briefly that it is difficult to understand everything without any prior knowledge. I found that most of my groupmates barely understood the topic even after the lectures, and needed my explanation and breakdown to understand at least some of it.

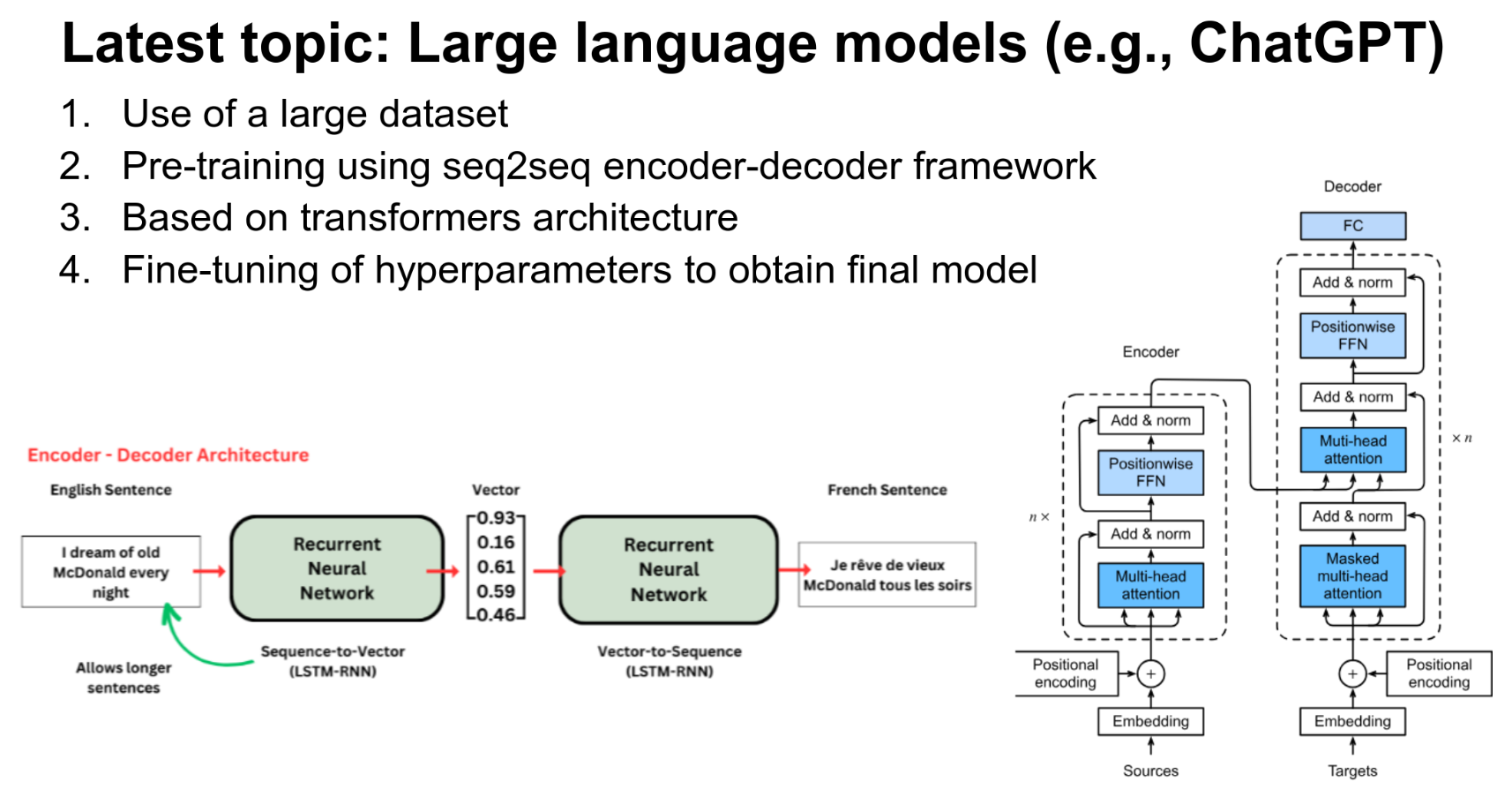


Figure 1: The slide talking about large language models.

Also, I think the professor teaching the technical portion of the AI section may not actually know what he is talking about, as he seems unsure of himself throughout the video. There are errors in the slide talking about large language models (Figure 1), as it seems to suggest that Generative Pre-trained Transformers (GPTs) are pre-trained using the seq2seq (sequence to sequence) encoder-decoder framework. This is incorrect, as GPTs are not pre-trained using seq2seq, but instead use the generative autoregressive target for pre-training, where the transformer continuously generates the next item (Chen et al., 2020). This is called generative pre-training (GP). The idea behind pre-training is to use a large dataset to inform the model of the characteristics of the dataset, and hopefully, the model will use this information to perform its intended task on the smaller dataset better through transfer learning. Seq2seq pre-training, while possible with a transformer, would be difficult to scale and hence not very suitable for pre-training since it requires an input and a corresponding output. This makes it difficult to pre-train a transformer on unlabelled data, such as data scraped from the internet, and increasing the efficacy of the model is extremely labour-intensive and time-consuming. Also, the transformer model that results from this seq2seq pre-training is not called a GPT, as it does not use generative pre-training (GP), but rather seq2seq pre-training. Seq2seq is more commonly used as the fine-tuning step, or the fourth step in the slide, for use cases like machine translation.

The diagrams shown in the slide (Figure 1), also do not show the GPT architecture. The diagram on the left describes a seq2seq LSTM-based autoencoder, which refers to the second point. The diagram on the right shows the original transformer model that makes use of both encoders and decoders, which refers to the third point. However, the GPT family of models are decoder-only models and hence do not have an encoder. Given that the points on the slide are specifically talking about GPTs, I would expect the diagrams to describe GPTs instead of other models. Also, large language models (LLMs) encompass a greater variety of models than just GPTs. Other LLMs include Mixture of Experts (MoE) used by Mixtral 8x7B, Mixtral 8x22B and Google Gemini 1.5, and encoder-only transformers, like BERT and XLNet.

Additionally, I found the panel videos to be very corporatised and impersonal. All of the panel experts just seemed to be promoting their company, and gave very generic answers that lacked the depth and the technicalities that they supposedly have as industry experts. A large part of the panel videos was just full of fluff and thinly veiled corporate advertising, and there were only a few small sections that contained useful information, new insights, or good advice. I expected far more depth and detail from a panel video with supposed industry experts.

# Impact on my interdisciplinary perspective

Truthfully, the course had little impact on my perspective, as I experienced the technical, social, business, governance and ethical impacts of technology being an open-source developer during the beginning of the current AI boom. It also helps that I completed CC0006 before this module, as it required the same consideration of these different perspectives for its assignments. However, I am glad to be working with my groupmates, as they suggested unique ideas I had not thought of and offered new and interesting insights. It is quite pleasant to have a group where my groupmates are willing to pull their weight and complete things instead of having to do almost everything myself like with regular CC modules. My groupmates having an opinion on the ideas I suggest is also a breath of fresh air compared to the usual where I’m coming up with everything and everyone else just agrees, mostly because they either don’t understand anything or just don’t care enough. The sharing during seminar sessions, maybe because the entire class is made up of scholars, is also insightful and worth listening to. This is absolutely not the case in regular CC modules, as people who want class participation points will just read out answers they have generated through ChatGPT, which sounds sophisticated and complex, but is really just hollow, full of fluff and lacks depth.

# Conclusion

Overall, my experience in this module can be summed up as average. There is nothing particularly exciting or mindblowing, and nothing particularly terrible.

# Bibliography

Chen, M., Radford, A., Child, R., Wu, J., Jun, H., Luan, D., Sutskever, I., Mark ChenOpenAI, S. F., Alec RadfordOpenAI, S. F., Rewon ChildOpenAI, S. F., Jeff WuOpenAI, S. F., Heewoo JunOpenAI, S. F., David LuanOpenAI, S. F., & Ilya SutskeverOpenAI, S. F. (2020, July 13). *Generative pretraining from pixels: Proceedings of the 37th International Conference on Machine Learning*. Guide Proceedings. https://dl.acm.org/doi/10.5555/3524938.3525096