References

This course drew from the following resources:

- <u>Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer</u> (Raffel et al, 2019)

- <u>Reformer: The Efficient Transformer</u> (Kitaev et al, 2020)

- <u>Attention Is All You Need</u> (Vaswani et al, 2017)

- <u>Deep contextualized word representations</u> (Peters et al, 2018)

- <u>The Illustrated Transformer</u> (Alammar, 2018)

- <u>The Illustrated GPT-2 (Visualizing Transformer Language Models)</u> (Alammar, 2019)

- <u>BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding</u> (Devlin et al, 2018)

- <u>How GPT3 Works - Visualizations and Animations</u> (Alammar, 2020)

Mark as completed