



References

This course drew from the following resources:

- [Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer](#) (Raffel et al, 2019)
- [Reformer: The Efficient Transformer](#) (Kitaev et al, 2020)
- [Attention Is All You Need](#) (Vaswani et al, 2017)
- [Deep contextualized word representations](#) (Peters et al, 2018)
- [The Illustrated Transformer](#) (Alammar, 2018)
- [The Illustrated GPT-2 \(Visualizing Transformer Language Models\)](#) (Alammar, 2019)
- [BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding](#) (Devlin et al, 2018)
- [How GPT3 Works - Visualizations and Animations](#) (Alammar, 2020)

Mark as completed