Transformers and Pretraining Module Suggested Readings

XCS224N Natural Language Processing with Deep Learning

1   Suggested Readings

1. Attention Is All You Need
2. The Illustrated Transformer
3. Transformer (Google AI blog post)
4. Layer Normalization
5. Image Transformer
6. Music Transformer: Generating music with long-term structure