1. What are Sequence-to-sequence models?
2. What are the Problem with Vanilla RNNs?
3. What is Gradient clipping?
4. Explain Attention mechanism
5. Explain Conditional random fields (CRFs)
6. Explain self-attention
7. What is Bahdanau Attention?
8. What is a Language Model?
9. What is Multi-Head Attention?
10. What is Bilingual Evaluation Understudy (BLEU)