Attention based Models and Transfer Learning

Assignment Questions

PW SKILLS

Attention based Models and Transfer Learning

1. What is BERT and how does it work?
2. What are the main advantages of using the attention mechanism in neural networks?
3. How does the self-attention mechanism differ from traditional attention mechanisms?
4. What is the role of the decoder in a Seq2Seq model?
5. What is the difference between GPT-2 and BERT models?
6. Why is the Transformer model considered more efficient than RNNs and LSTMs?
7. Explain how the attention mechanism works in a Transformer model.
8. What is the difference between an encoder and a decoder in a Seq2Seq model?
9. What is the primary purpose of using the self-attention mechanism in transformers?
10. How does the GPT-2 model generate text?
11. What is the main difference between the encoder-decoder architecture and a simple neural network?
12. Explain the concept of “fine-tuning” in BERT.
13. How does the attention mechanism handle long-range dependencies in sequences?
14. What is the core principle behind the Transformer architecture?
15. What is the role of the "position encoding" in a Transformer model?
16. How do Transformers use multiple layers of attention?
17. What does it mean when a model is described as “autoregressive” like GPT-2?
18. How does BERT's bidirectional training improve its performance?
19. What are the advantages of using the Transformer over RNN-based models in NLP?
20. What is the attention mechanism’s impact on the performance of models like BERT and GPT-2?

Practical

1. How to implement a simple text classification model using LSTM in Keras?
2. How to generate sequences of text using a Recurrent Neural Network (RNN)?
3. How to perform sentiment analysis using a simple CNN model?
4. How to perform Named Entity Recognition (NER) using spaCy?
5. How to implement a simple Seq2Seq model for machine translation using LSTM in Keras?
6. How to generate text using a pre-trained transformer model (GPT-2)?
7. How to apply data augmentation for text in NLP?
8. How can you add an Attention Mechanism to a Seq2Seq model?