

Summary and Highlights

Congratulations! You have completed this lesson. At this point in the course, you know that:

In the decoder method, the target sequence and the memory are taken as inputs, and the target sequence receives token embedding and positional encoding, like the source sequence.

The transformer layer handles both encoding and decoding processes.

Transformers process all text sequences simultaneously for language translation.

Cross-attention computes attention scores between each target position and all source positions.