Encoder-Decoder Architecture

Encoder

Converts input sequence to context vector

(Typically LSTM/GRU)

→

Context Vector

Fixed-size representation of entire input

Decoder

Generates output conditioned on context

(Typically LSTM/GRU)

→

🎓 **Training**

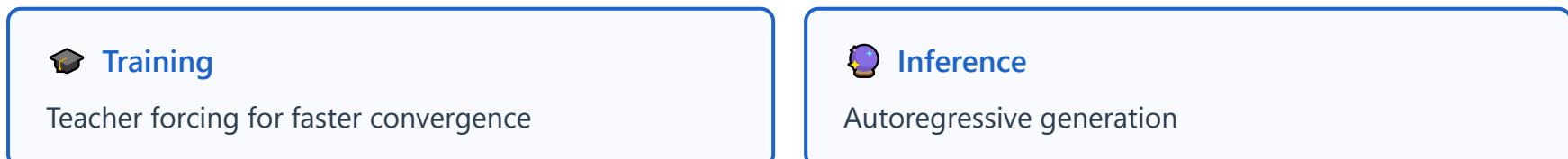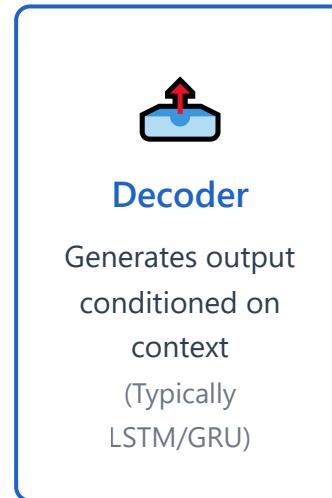Teacher forcing for faster convergence

🔮 **Inference**

Autoregressive generation

Both encoder and decoder typically use **LSTM** or **GRU** cells