

BERT vs GPT: Different Objectives, Different Architectures

BERT: Bidirectional Encoder

GPT: Unidirectional Decoder

Can see both left AND right context

Can only see left context (for generation)

