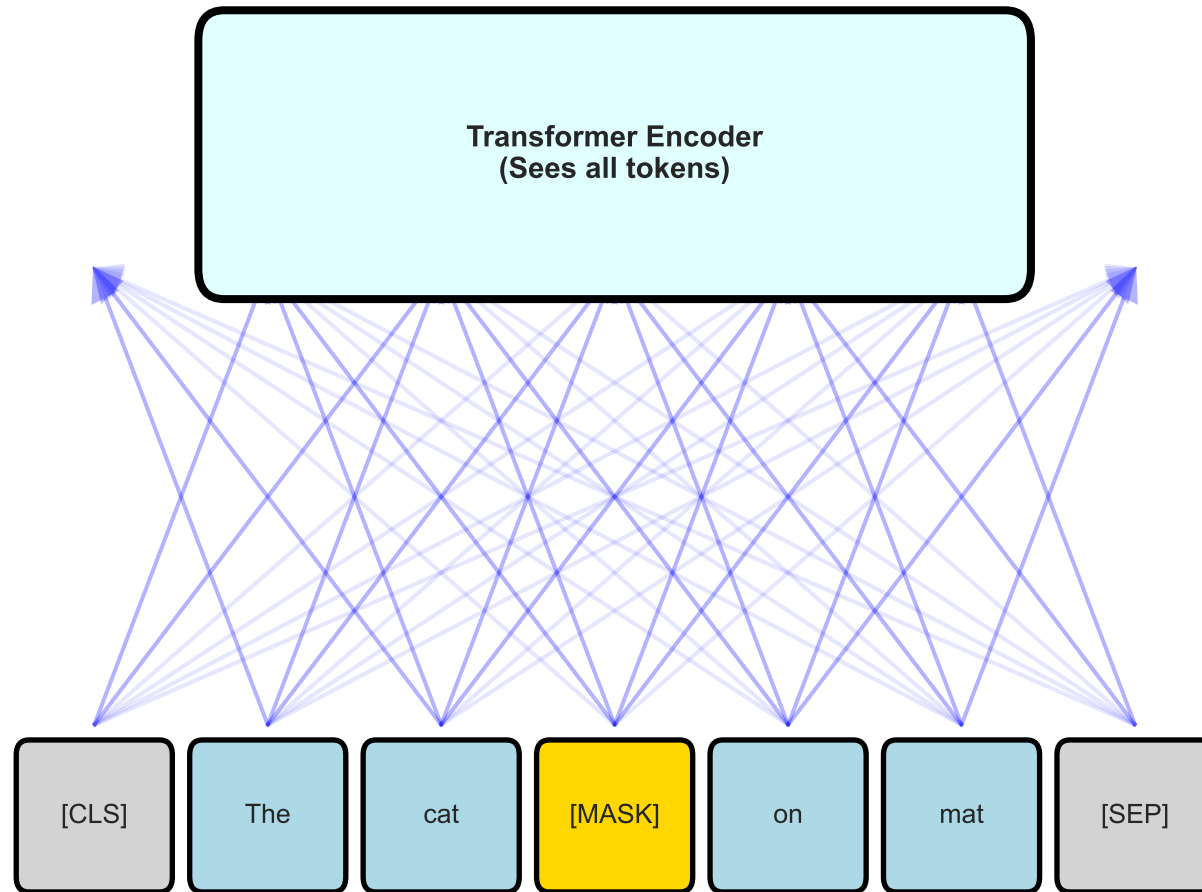


# BERT vs GPT: Different Objectives, Different Architectures

## BERT: Bidirectional Encoder

Can see both left AND right context



## GPT: Unidirectional Decoder

Can only see left context (for generation)

