

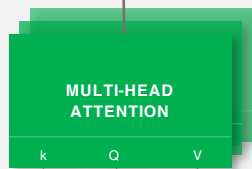
# ENCODER $N \times$

POSITIONAL  
ENCODING



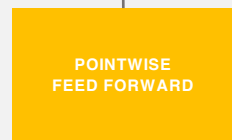
EMBEDDING

INPUTS



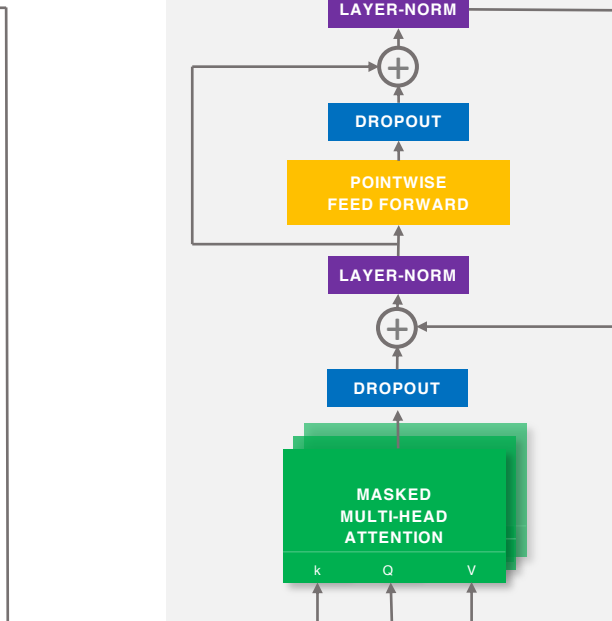
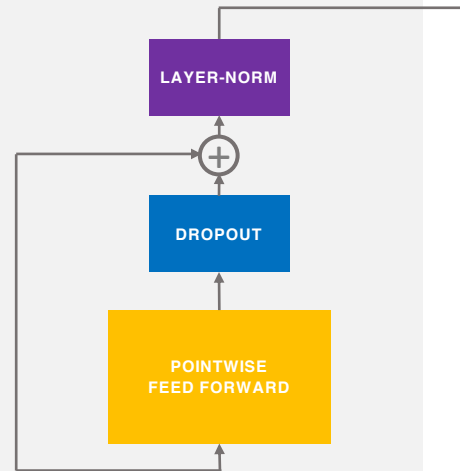
DROPOUT

LAYER-NORM



DROPOUT

LAYER-NORM



# DECODER $N \times$

POSITIONAL  
ENCODING



EMBEDDING

PREV. OUTPUTS



DROPOUT

LAYER-NORM



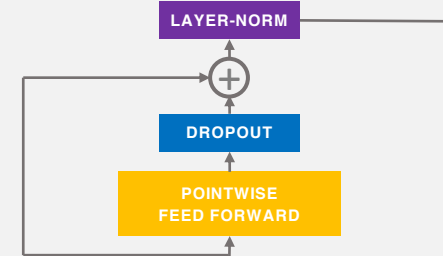
DROPOUT

LAYER-NORM



DROPOUT

LAYER-NORM



OUTPUT  
PROBABILITIES

