

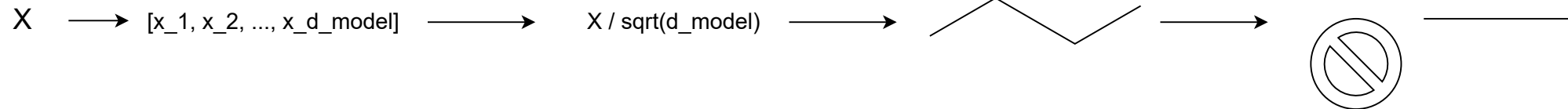
Encoder

Word Embedding

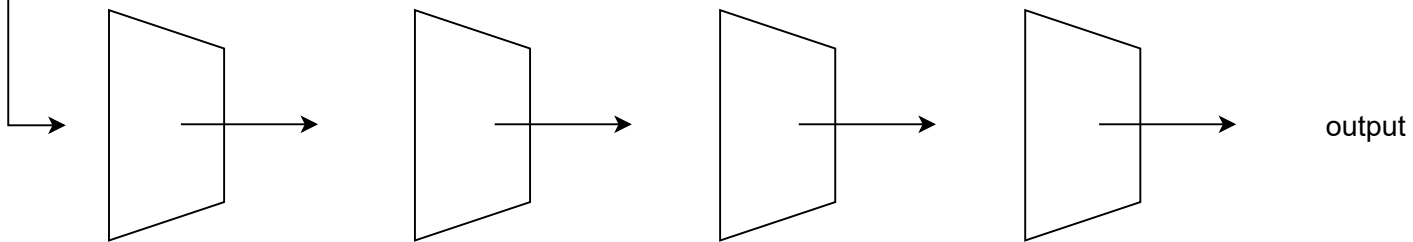
Scaling

Positional Encoding

dropout
0.1



Encoder Layers



$N = 4$