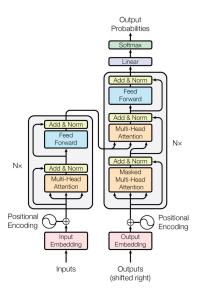
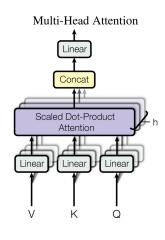
Understand the hyper parameters book books ize ength block-size hello $\begin{bmatrix} +1.2 \\ -1.8 \end{bmatrix}$ dim : dim \longrightarrow n-embed. self-attention Block This has number of needs in each Block n-largers A Mumber of Blocks Number of layers



Transformes architedure



multi- read attention.