query — key — value

$\bar{x}_1$ → $\bar{q}_1$ $\bar{k}_1$ $\bar{v}_1$

$\bar{x}_k$ → $\bar{q}_k$ $\bar{k}_k$ $\bar{v}_k$

FF

Self-att

ene

$$a_{ij} = \underset{j}{\text{softmax}}\left(\frac{1}{\sqrt{d}}\, \bar{q}_i^T \bar{k}_j\right)$$

$$\bar{z}_i = \sum_{i=1}^{K} a_{ij}\, \bar{v}_j$$

$\bar{x}_i \in \mathbb{R}^D$

$\bar{q}_i = W_1^Q \bar{x}_i \in \mathbb{R}^d$

$\bar{k}_i = W_1^K \bar{x}_i \in \mathbb{R}^d$

$\bar{v}_i = W_1^V \bar{x}_i \in \mathbb{R}^{d'}$

X

$Q_1 = W_1^Q X$ $D \times d$

$K_1 = W_1^K X$

$V_1 = W_1^V X$ $D \times d'$

$$Z_1 = \text{softmax}\left(\frac{1}{\sqrt{d}} Q_1 K_1^T\right) \cdot V_1 \in \mathbb{N} \times d'$$

$W_0^Q, W_1^Q, \dots, W_7^Q$ ---- $W_7^V = S-1$

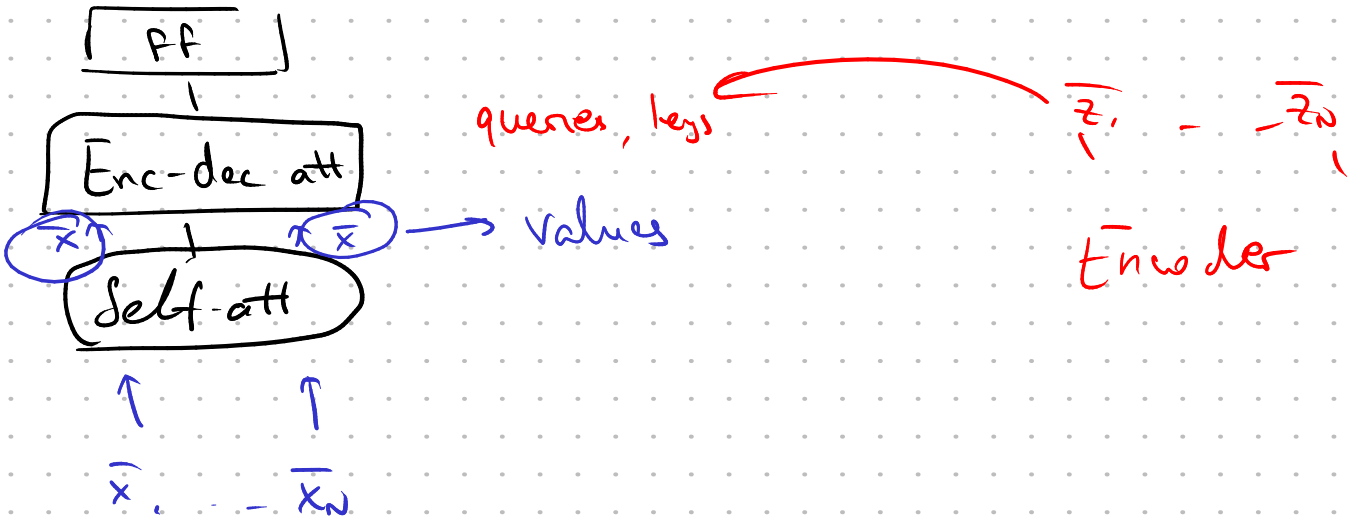
S модули

$D \times d, D \times d, D \times d'$ — " — W_7^K

+1 модуль $(Ns \times N)$

X → Z_0 | Z_{S-1} → $N \boxed{Z_0 Z_1 \dots Z_{S-1}}$ Nd' · $\boxed{W^0}$ $Nd' \times D$ → Z $N \times D$

$N \times d'$


FF
Enc-dec att
Self-att
queries, keys
values
Encoder