

Perceptron: General Perceptron with Static Attention

Self-Attention

$$Z_{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V ; Q \in \mathbb{R}^{m \times d}, K \in \mathbb{R}^{m \times d}, V \in \mathbb{R}^{m \times c} \Rightarrow O(M^2)$$

• N - Sequence length

• If $k \neq v$ derived by input array, let

Q be derived by a latent array with index $\dim(N \times L \times M)$, the complexity becomes $O(MN)$

(Refers to pages 3-12)

• (kx) Fourier feature positional encoding;

i) directly represents the positional structure of the input data

(preserving 2D temporal or

2D spatial structure)

ii) control the number of frequency bands in our positional encoding independently of the cutoff frequency

iii) log-uniformly sample all frequencies up to a target resolution

Cross-Att Module

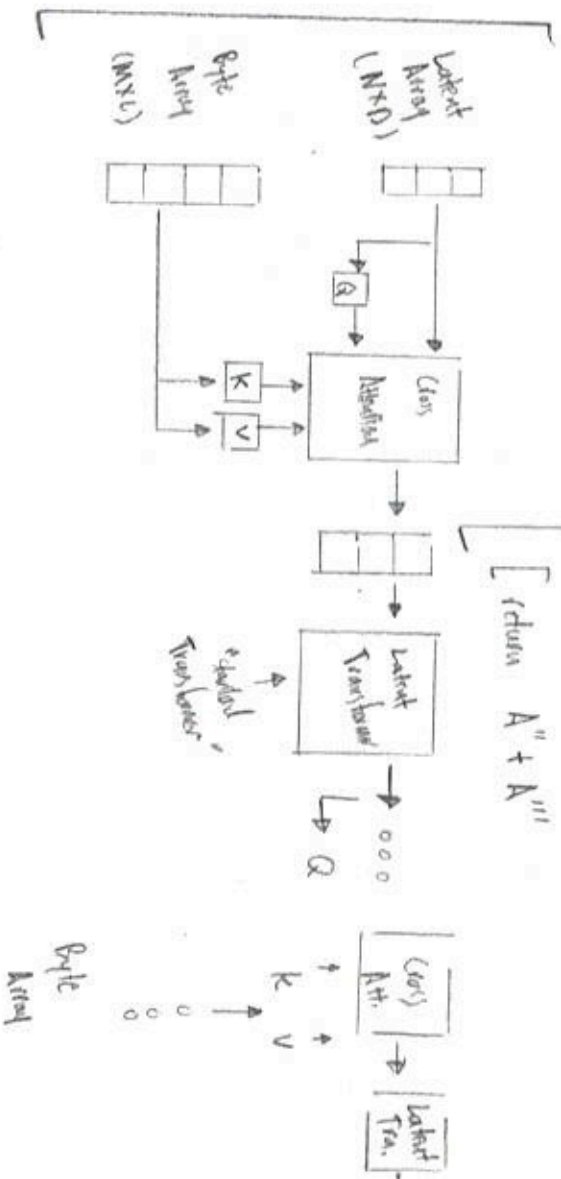
[Latent \rightarrow Layer Norm] $\rightarrow L'$
[Byte \rightarrow Layer Norm] $\rightarrow B'$

[Dense (project Dim) (L')] $\rightarrow Q$
[Dense (project Dim) (B')] $\rightarrow K$
[Dense (project Dim) (B')] $\rightarrow V$

[Attention (Q, K, V)] $\rightarrow A$
[$A + \text{Latent}$] $\rightarrow A'$

[$A' \rightarrow$ Layer Norm] $\rightarrow A''$
[Dense (the dim) (A'')] $\rightarrow A'''$

[return $A'' + A'''$]



"Shares weights of cross-att module, Latent-Tra, & latent array, forming a RNN"

Perceptron: Dim Sanity Check

M sequence length

Byte Array $\rightarrow B_{M \times C}$

Latent Array $\rightarrow L_{N \times P}$

Projection Dim $\rightarrow P$

Cross-Attention Module

$$\left[\begin{array}{l} L_{N \times P} W_{Q \times P}^Q \rightarrow Q_{N \times P} \\ B_{M \times C} W_{C \times P}^K \rightarrow K_{M \times P} \\ B_{M \times C} W_{C \times P}^V \rightarrow V_{M \times P} \end{array} \right] \rightarrow \text{softmax} \left(\frac{Q_{N \times P} K_{M \times P}^T}{\sqrt{d_Q}} \right) V_{M \times P} \rightarrow A_{N \times P}$$

"Same dim as original latent $L_{N \times P}$ "

Latent Transformer (normal transformer unit where "latent" output from cross-att is used)

