‹¡{—lèa R›wé–5ÿ Attention Matrixÿ f/Transformerj!W‹N-v„h8_Ãkeš¤ÿ Qvvîv„f/• •Ç‹¡{—•“Qe^•R N-T N*C

**1. •“QeQÆY **
PG‹¾•“Qe^•R N: \(X \in \mathbb{R}^{N \times C} \)ÿ QvN- \(N \) f/^•R ••^¦ÿ tokenep‘Ïÿ ÿ \(C \) f/kÏN*to
- **gå‹âwé–5ÿ Queryÿ **ÿ \(Q = X W_Q \)
- **•.wé–5ÿ Keyÿ **ÿ \(K = X W_K \)
- **P<wé–5ÿ Valueÿ **ÿ \(V = X W_V \)

QvN- \(W_Q, W_K, W_V \in \mathbb{R}^{C \times d} \) f/Sï[fN`SÂepÿ \(d \) f/lèa R›Y4v„~ô^¦0

---

2. ‹¡{—g*_RN S v„lèa R›R ep
• •Çwé–5NXlÕ‹¡{—b@g token[ùNK•ôv„vøQs`'R epÿ
\[
P = Q K^T \in \mathbb{R}^{N \times N}
\]
kÏN*QC} \(P_{ij} \) ˆhy:{, \(i \) N*token[ù{, \(j \) N*tokenv„Qslèz ^¦0

**3. •)e>ÿ Scalingÿ **
N:–2kbp¹yïP<•ÇY'[ü•ôh¯^¦mˆY1ÿ [ùR ep•ÛˆL•)e>ÿ –dNå \(\sqrt{d} \)ÿ ÿ
\[
P_{\text{scaled}} = \frac{P}{\sqrt{d}}
\]

**4. c©x ÿ Sï• ÿ **
W(‰ãx VhN-ÿ N:••QMg*geOá`olÄ—2ÿ O O•u(N N ‰Òc©x ÿ masked attentionÿ ÿ
\[
P_{\text{masked}} = P_{\text{scaled}} + M
\]
QvN- \(M \) f/c©x wé–5ÿ N N ‰ÒN: \(-\infty \)ÿ QvOYN:0ÿ 0

**5. Softmax_RN S **
[ùkÏN ˆLÿ kÏN*tokenv„lèa R›R ^ ÿ ^”u(Softmaxÿ _—R0_RN S v„lèa R›gC‘Íwé–5 \(A \)ÿ
\[
A = \text{Softmax}(P_{\text{scaled}}) \in \mathbb{R}^{N \times N}
\]
SoftmaxxnOÝkÏˆLgC‘ÍTŒN:1ÿ N —^• 0

---

6. R gClBTŒ
u(lèa R›gC‘Í[ùP<wé–5 \(V \) R gClBTŒÿ _—R0•“Qúÿ
\[
\text{Output} = A V \in \mathbb{R}^{N \times d}
\]

Ql_ `;~Ó
[Œetv„lèa R›wé–5‹¡{—Sïˆhy:N:ÿ
\[
A = \text{Softmax}\left(\frac{Q K^T}{\sqrt{d}}\right)
\]

**Nãx y:O‹ÿ PyTorchÿ **
```python
import torch
import torch.nn.functional as F

def compute_attention(Q, K, V, mask=None):
    d_k = Q.size(-1)
    scores = torch.matmul(Q, K.transpose(-2, -1)) / torch.sqrt(torch.tensor(d_k))
    if mask is not None:
        scores = scores.masked_fill(mask == 0, -1e9)
    attn_weights = F.softmax(scores, dim=-1)
    output = torch.matmul(attn_weights, V)
    return output, attn_weights
```

R¨` z u•S ÿ DynamicViTÿ v„e9•Û
W(DynamicViTN-ÿ • •Ç˜„mKj!WWu b NŒP<c©x \(\hat{D} \)ÿ ^vOîe9lèa R›‹¡{—ÿ ‰Á‹ºe‡Ql_ 10-11ÿ ÿ
1. **c©x g„• **ÿ OÝuY•ê•Þc¥ÿ \(G_{ii}=1 \)ÿ ÿ —^•ê•Þc¥•èR u1 \(\hat{D} \) Q³[š0
2. **c©x ^”u(**ÿ W(SoftmaxRM\ eàQstokenv„lèa R›R ep•n–öÿ [žs°R¨` Rjg•0

•ÙyÍe9•ÛQÏ\ N†Q—OY‹¡{—ÿ T eöOÝc N†xlNöSËY}v„z u•`'0