$x_1 \quad x_2 \quad \cdots \quad x_n$

bag of words

RNN

Position Encoding

Absolute Position → trainable embeddings

$(e_i+p_i)^T W_Q^T W_k (e_j+p_j)$

$$\begin{cases} P_{(k,2i)} = \sin(k/10000^{2i/d}) \\ P_{(k,2i+1)} = \cos(k/10000^{2i/d}) \end{cases}$$

$e_i^T W_Q^T W_k e_j + p_i^T W_Q^T W_k e_j + e_i^T W_Q^T W_k p_j + p_i^T W_Q^T W_k p_j$

$p_i^T W_Q^T W_k p_j = g(i-j)$

$p_i^T p_j = g'(i-j)$

$\cos\theta\cos\varphi + \sin\theta\sin\varphi = \cos(\theta-\varphi)$

$\cos[(i-j)\theta_0]$

$\theta_k = k\theta_0$

$\cos(i\theta_0 - j\theta_0)$

$p_i = \begin{pmatrix} \cos\theta_i \\ \sin\theta_i \end{pmatrix} \quad p_j = \begin{pmatrix} \cos\theta_j \\ \sin\theta_j \end{pmatrix}$

$\cos(\theta_i - \theta_j)$

Relative Position.

$(e_i+u)^T W_Q^T W_k (e_j+r_{i-j})$

XLNet. ① $+ u^T W_Q^T W_k e_j + e_i^T W_Q^T W_k r_{i,j} + v^T W_Q^T W_k r_{i,j}$

$(E+P)^T W_Q^T W_k (E+P)$

TS ① $+ p_i^T W_Q^T W_k p_j \Rightarrow \beta_{i,j}$

DeBERTa ① $+ r_{j,i}^T W_Q^T W_k e_j + e_i^T W_Q^T W_k r_{ij}$

Rotate Position → $f(\vec{q},m)^T f(\vec{k},n) = g(\vec{q},\vec{k},m-n)$

↓

ViT

$x_i + p_i$? $f(\cdot)$

$$f(\vec{q},m)^T f(\vec{k},n) = g(\vec{q},\vec{k},m-n)$$

$$f(\vec{q},m) = \begin{bmatrix} \cos m\theta & -\sin m\theta \\ \sin m\theta & \cos m\theta \end{bmatrix} \vec{q}$$

$$f(\vec{q},m)^T f(\vec{k},n) = \vec{q}^T \begin{bmatrix} \cos m\theta & -\sin m\theta \\ \sin m\theta & \cos m\theta \end{bmatrix}^T \begin{bmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{bmatrix} \vec{k}$$

$$= \vec{q}^T \begin{bmatrix} \cos(m-n)\theta & \sin(m-n)\theta \\ -\sin(m-n)\theta & \cos(m-n)\theta \end{bmatrix} \vec{k}$$

$$= g(\vec{q},\vec{k},m-n)$$

标题	年
BERT: Pre-training of Deep Bidirectional Transformers for L...	2019
Transformer升级之路：4、二维位置的旋转式位置编码 - 科学空...	
Transformer升级之路：2、博采众长的旋转式位置编码 - 科学空...	
让研究人员绞尽脑汁的Transformer位置编码 - 科学空间|Scientifi...	
Transformer升级之路：1、Sinusoidal位置编码追根溯源 - 科学...	
RoFormer: Enhanced Transformer with Rotary Position Emb...	2021
Transformer-XL: Attentive Language Models Beyond a Fixed...	2019
Attention Is All You Need	2017