

Vaswani et al. (2017) provide sine and cosine functions so that we can generate different frequencies for the positional encoding (PE) for each position and each dimension i of the $d_{model} = 512$ of the word embedding vector:

$$\begin{aligned} PE_{(pos\ 2i)} &= \sin\left(\frac{pos}{10000^{\frac{2i}{d_{model}}}}\right) \\ PE_{(pos\ 2i+1)} &= \cos\left(\frac{pos}{10000^{\frac{2i}{d_{model}}}}\right) \end{aligned} \tag{1}$$