**位置编码**

为什么需要位置编码？

答：位置编码体现token之间的时序/顺序关系；只有对位置敏感的模型才需要添加额外的位置编码，例如RNN、TextCNN。

两种编码方式——绝对位置编码 or 相对位置编码

答：

绝对位置编码：直接对不同的位置随机初始化一个postion embedding，加上token\_embedding就行。体现的是不同位置编码一定不同。

缺陷：位置之间没有约束条件，但有时位置之间是有一定的距离

相对位置编码：

1. Sinusoidal Position Encoding

采用sin-cos的方式，能学到位置的相对关系，主要体现在正弦余弦公式上；但是无法区分方向。

1. Complex embedding

对比：

1. 从方法的可理解性上，相比相对位置编码的两种方法，Learned Positional Embedding更加的简单直接，易于理解。
2. 从参数维度上，使用Sinusoidal Position Encoding不会引入额外的参数，Learned Positional Embedding增加的参数量会随max\_seq\_len线性增长，而Complex Embedding在不做优化的情况下，会增加三倍word embedding的参数量。
3. 在可扩展性上，Learned Positional Embedding可扩展性较差，只能表征在max\_seq\_len以内的位置，而另外两种方法没有这样的限制，可扩展性更强。