

## NLP方向总结-注意力机制

### attention机制

Self Attention顾名思义，指的不是Target和Source之间的Attention机制，而是Source内部元素之间或者Target内部元素之间发生的Attention机制，也可以理解为Target=Source这种特殊情况下的注意力计算机制，Self Attention后会更容易捕获句子中长距离的相互依赖的特征

soft/hard attention, Soft Attention是所有的数据都会注意，都会计算出相应的注意力权值，不会设置筛选条件  
Hard Attention会在生成注意力权重后筛选掉一部分不符合条件的注意力，让它的注意力权值为0，即可以理解为不再注意这些不符合条件的部分。

global/local attention  
global attention: 传统attention，对所有encoder输出进行计算  
local attention: 介于soft和hard之间，会预测一个位置并选取一个窗口进行计算