

## 论文: Relation Classification via Convolutional Deep Neural Network

模型的输入特征:

本论文使用的模型含有一些精心设计的特征:

**lexical 特征:** 两个目标词的词向量 (L1, L2), 两个目标词相邻单词的词向量 (L3, L4), 目标词的上位词 (WordNet hypernyms L5); 如下图所示:

Features	Remark
L1	Noun 1
L2	Noun 2
L3	Left and right tokens of noun 1
L4	Left and right tokens of noun 2
L5	WordNet hypernyms of nouns

Table 1: Lexical level features.

**sentence 特征:** 本单词的词向量, 相邻单词的词向量, 到目标词之间的距离

词向量特征:  $WF = \{[x_s, x_0, x_1], [x_0, x_1, x_2], \dots, [x_5, x_6, x_e]\}$

距离特征:  $PF = [d1, d2]$

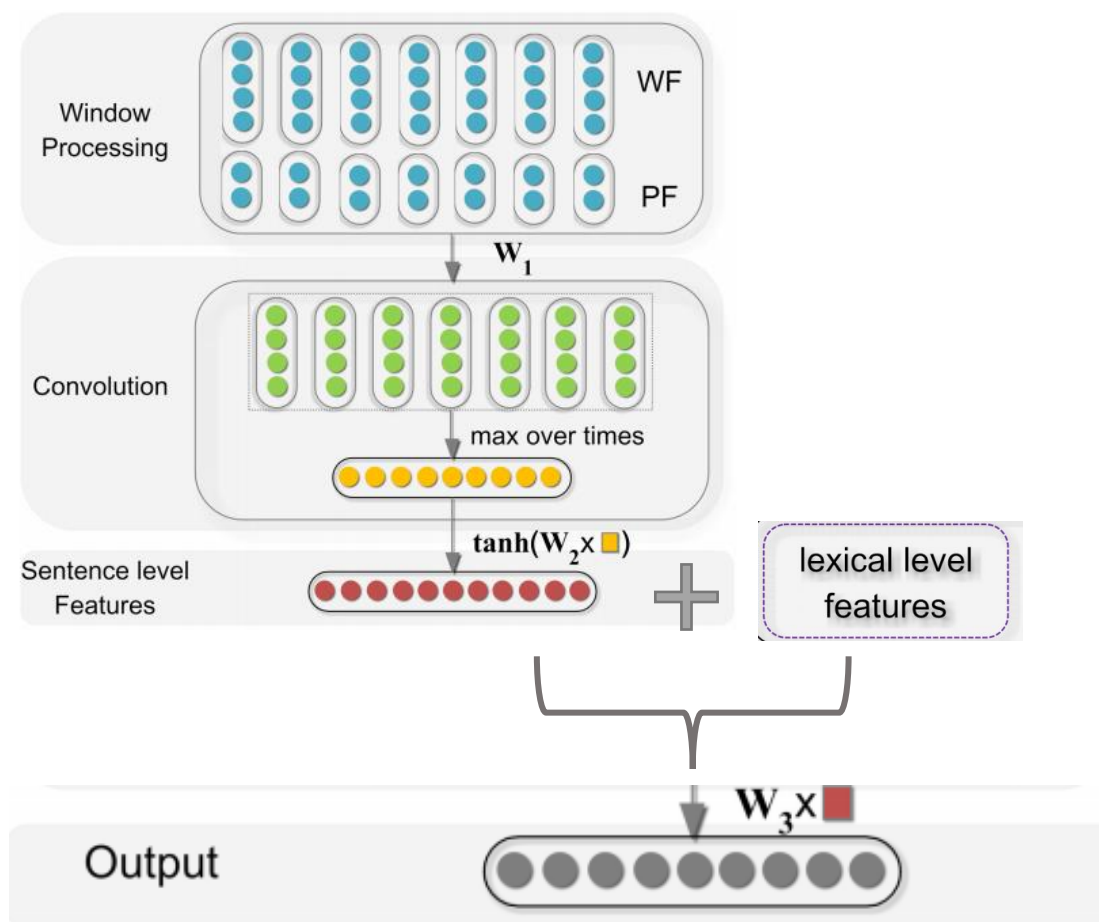
拼接后输出:  $[WF, PF]$

The Effect of Learned Features:

	Feature Sets	F1
Lexical	L1	34.7
	+L2	53.1
	+L3	59.4
	+L4	65.9
	+L5	73.3
Sentence	WF	69.7
	+PF	78.9
Combination	all	82.7

根据论文的结论, 文中使用的 features 对最终的 F1 均有较大的提升。

模型较为直观：



1. 将所有的 **sentence 特征** 拼接，与矩阵  $W_1$  相乘做一个线性变换（无激活函数），再使用一个 global maxpooling 的操作（这里就是标题中 Convolutional）；
2. 将 maxpooling 的结果传入  $W_2$  全连接网络（ $\tanh$  激活），输出记为 sentence feature
3. **拼接** sentence feature 和 lexical feature，最后接上输出层  $W_3$ ，使用 softmax 输出结果。