词性标注实现方法

# 神经网络算法

## 神经网络参数

①、网络节点

输入层神经元个数由输入词组的个数确定，125\*3=375，即输入由待标记次的前后各两个词的词性概率及待标记次本身的词性概率表达。

输出层神经元个数由词的标注集大小确定，标注集大小为125，故输出层神经元个数为125。

隐含层神经元个数，经验公式s=sqrt(0.43mn+0.12nn+2.54m+0.77n+0.35)+0.51，其中m为输入层个数，n为输出层个数；s = 152，但是由于神经网络过大，训练时间过长，在试验后修改为100

②、初始权值的确定

初始权值是不应完全相等的一组值。已经证明，即便确定  存在一组互不相等的使系统误差更小的权值，如果所设*W*ji的的初始值彼此相等，它们将在学习过程中始终保持相等。故而，在程序中，我们设计了一个随机发生器程序，产生一组一0.5~+0.5的随机数，作为网络的初始权值。

③、最小训练速率  在经典的BP算法中，训练速率是由经验确定，训练速率越大，权重变化越大，收敛越快；但训练速率过大，会引起系统的振荡，因此，训练速率在不导致振荡前提下，越大越好。因此，在DPS中，训练速率会自动调整，并尽可能取大一些的值，但用户可规定一个最小训练速率。该值一般取0.9。

④、动态参数  动态系数的选择也是经验性的，一般取0.6 ~0.8。

⑤、Sigmoid参数 该参数调整神经元激励函数形式，一般取0.9~1.0之间。