

# Crf

1. 条件随机场用途：求序列的隐状态
2. 条件随机场思想：

$$\text{score}(l|s) = \sum_{j=1}^m \sum_{i=1}^n \lambda_j f_j(s, i, l_i, l_{i-1})$$

句子  $s$  (就是我们要标注词性的句子)

$i$ , 用来表示句子  $s$  中第  $i$  个单词

$l_i$ , 表示要评分的标注序列给第  $i$  个单词标注的词性

$l_{i-1}$ , 表示要评分的标注序列给第  $i-1$  个单词标注的词性

它的输出值是 0 或者 1, 0 表示要评分的标注序列不符合这个特征, 1 表示要评分的标注序列符合这个特征。

线性链 crf:

特征函数仅仅依靠当前单词的标签和它前面的单词的标签对标注序列进行评判, 这是 CRF 中的一种简单情况。

也可以用非线性链条件随机场:

针对上面公式举例子:

1) 虽然  $\text{score}(l|s)$  的函数的自变量包含 4 个, 但是真正用的时候并不是四个都要用。

如: 当  $l_i$  是“副词”并且第  $i$  个单词以“ly”结尾时, 我们就让  $f_1 = 1$ , 其他情况  $f_1$  为 0。不难想到,  $f_1$  特征函数的权重  $\lambda_1$  应当是正的。而且  $\lambda_1$  越大, 表示我们越倾向于采用那些把以“ly”结尾的单词标注为“副词”的标注序列。

这个例子就仅仅使用了  $l_i$  和  $i$  和  $s$  这三个变量, 和前面一个词对应的隐变量没关系

2)  $i=1$ ,  $l_i$ =动词, 并且句子  $s$  是以“?”结尾时,  $f_2=1$ , 其他情况  $f_2=0$ 。同样,  $\lambda_2$  应当是正的, 并且  $\lambda_2$  越大, 表示我们越倾向于采用那些把问句的第一个单词标注为“动词”的标注序列。

这个例子就是用了两个变量,  $s$  和  $l_i$

3) 当  $l_{i-1}$  是介词,  $l_i$  是名词时,  $f_3 = 1$ , 其他情况  $f_3=0$ 。  $\lambda_3$  也应当是正的, 并且  $\lambda_3$  越大, 说明我们越认为介词后面应当跟一个名词。

Emmm.....

3. 和 hmm 的比较

hmm 是 crf 的一种特殊情况。Crf 远比 hmm 要强大。

Crf 可以做到, 但是 hmm 做不到的例子:

CRF 可以定义数量更多, 种类更丰富的特征函数。HMM 模型具有天然具有局部性, 就是说, 在 HMM 模型中, 当前的单词只依赖于当前的标签, 当前的标签只依赖于前一个标签。这样的局部性限制了 HMM 只能定义相应类型的特征函数。

譬如上面的:  $i=1$ ,  $l_i$ =动词, 并且句子  $s$  是以“?”结尾时,  $f_2=1$ , 其他情况  $f_2=0$ 。

这个例子 hmm 就做不到。但是 crf 可以做到。

Crf 转化为 hmm 的一个例子:

Hmm:

$p(l_i|l_{i-1})$  是转移概率, 比如,  $l_{i-1}$  是介词,  $l_i$  是名词, 此时的  $p$  表示介词后面的词是名词的概率。

$p(w_i|l_i)$ 表示发射概率，比如  $l_i$  是名词， $w_i$  是单词“ball”，此时的  $p$  表示在是名词的状态下，是单词“ball”的概率。

$$p(l, s) = p(l_1) \prod_i p(l_i | l_{i-1}) p(w_i | l_i)$$

给  $\text{hmm}$  加  $\log$

$$\log p(l, s) = \log p(l_0) + \sum_i \log p(l_i | l_{i-1}) + \sum_i \log p(w_i | l_i).$$

下面是  $\text{crf}$ :

$$\text{score}(l|s) = \sum_{j=1}^m \sum_{i=1}^n \lambda_j f_j(s, i, l_i, l_{i-1})$$

做比较可知:

对于  $\text{HMM}$  中的每一个转移概率  $p(l_i=y|l_{i-1}=x)$ , 我们可以定义这样的一个特征函数:

$$f_{x,y}(s, i, l_i, l_{i-1}) = 1$$

同样的，对于  $\text{HMM}$  中的每一个发射概率，我们也都可以定义相应的特征函数，并让该特征函数的权重等于  $\text{HMM}$  中的  $\log$  形式的发射概率。

#### 4. 和逻辑回归的比较

$$p(l|s) = \frac{\exp[\text{score}(l|s)]}{\sum_{l'} \exp[\text{score}(l'|s)]} = \frac{\exp[\sum_{j=1}^m \sum_{i=1}^n \lambda_j f_j(s, i, l_i, l_{i-1})]}{\sum_{l'} \exp[\sum_{j=1}^m \sum_{i=1}^n \lambda_j f_j(s, i, l'_i, l'_{i-1})]}$$

逻辑回归是用于分类的对数线性模型，条件随机场是用于序列化标注的对数线性模型

#### 5. 条件随机场参数求解方法

就是求解逻辑回归的参数的方法。

#### 6. 条件随机场如何解码