

流程:

1. 训练 coarse 模型, 得到 aspect mention matrix.
 $shape = (aspects\ number, vector\ number, vector\ dim)$

2. 将 aspect matrix reshape 成为一个新的 matrix.

$shape = [1, aspects\ num * vector\ number, vector\ dim]$

3. 构建一个中间模型, 该模型的 aspect matrix 由 step 2 的 new matrix 初始化.

4. 用中间模型 (media model) 计算出每个 fine grained attribute 包含的每个 instance 对 new matrix 中每个 vector 的 attention.

5. 这样, 对每个 ^{fine grained} attribute, 可以得到一个 attention matrix:

$$A_j = \begin{bmatrix} a_{j1} & a_{j2} & \dots & a_{ji} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \end{bmatrix} \quad \text{句子的 attention vector.}$$

a_{ji}^k 表示 fine grained attribute j 的第 k 个 instance 对 new matrix 中 vector i 的 attention

6. $numpy.mean(A_j, axis=0)$.

由此得到 fine grained attribute j 对每个 new matrix 中 vector 的比重. 即

$$A'_j = [a'_{j1} \ a'_{j2} \ \dots \ a'_{jd}]$$

所以, 所有 A'_j 可形成一个新矩阵:

$$W = \begin{bmatrix} A'_1 \\ A'_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} a'_{11} & a'_{12} & \dots & a'_{1d} \\ a'_{21} & \dots & & \\ \vdots & & & \vdots \end{bmatrix}$$

a'_{ji} 表示 attribute j 对 vector i 的 attention

7. 用 $x_{ij} \in \{0, 1\}$ 表示是否用 vector i 去初始化 fine grained model 中 attribute j 的某个 vector.
那么, 我们求:

$$\max \sum_{j=1}^{|A|} \sum_{i=1}^d x_{ij} w_{ij}$$

$$\text{constrains: } \sum_{j=1}^{|A|} x_{ij} \leq 2 \text{ or } 3.$$

$$\sum_{i=1}^d x_{ij} = 3.$$

此为一个 integer linear programming 问题.
这样做的目的是, 使 attention 之和最大.

原理:

Coarse 模型训练出 Aspect matrix

$$M = \begin{bmatrix} x_{11} & x_{12} & \dots \\ \vdots & \vdots & \\ x_{n1} & x_{n2} & \dots \end{bmatrix} \text{ vector } l.$$

对每个 attribute A_j , 它包含的句子构成的矩阵:

$$S = \begin{bmatrix} s_{11} & s_{12} & \dots \\ \vdots & \vdots & \\ \vdots & \vdots & \end{bmatrix} \text{ 句子 } l.$$

每个句子 s_i 可以对每个 vector $i \in M$ 算出一个 attention

那么, 由此可得到 A_j 对每个 vector 的 attention.

因此, 针对每个 A_j , 应挑选 attention 最大的 k 个 vector. 当考虑所有 attribute 时, 应该使 attention 之和最大 (在满足 constraints 的同时).