**Reference Type:**  Journal Article

**Record Number:** 35

**Author:** Yukun Ma, ∗ Haiyun Peng, † Erik Cambria†

**Year:** 2018(AAAI)

**Title:** <Targeted Aspect-Based Sentiment Analysis via Embedding.pdf>

**Short Title:** <Targeted Aspect-Based Sentiment Analysis via Embedding.pdf>

**'File' Attachments:** internal-pdf://0308681767/Targeted Aspect-Based Sentiment Analysis via E.pdf

1. 论文综述及分析

* 本文解决的问题是“分析人对特定属性的意见和情感极性”。
* 本文提出的东西有：层次attention机制（即两层attention：target-level和sentence-level）、在LSTM中引入外部知识的两种方法（修改LSTM的结构和添加在输入中）。效果超过了targeted aspect sentiment tasks上state-of-the-art 的方法。
* 对于存在多个target的attention问题，当前的解决方法是：求多个target的和平均向量或者直接忽视这个问题，而本文提出的是“weights each target word with an attention weight”，即每个target满足“target is represented by its most informative components”，这是target-level的attention，这是从单词或短语方面的考虑。
* Sentence-level的attention是从全局的考虑，并且向量中的每个元素都编入了aspect 和target对的信息，现存的研究一般都把aspect和target用于查询。
* 知识的使用，利用外部知识生成隐藏层输出，并且控制信息的流动（跟遗忘门作用相似）

1. 详细分析

* targeted aspect-based sentiment analysis（ABSA）
* target是一个单词（Brent）或短语（west London），并且规定指代同一个target的不同表述都用同一个target术语表示，所以每个target都可以表示为T = {t1, t2, … , tm}，ti表示target中第i个词在sentence中的位置，如“I live in [West London] for years”，TWest London = {3,4}。
* 层级attention机制（LSTM外）
* Target-level的attention建立在encoder的隐藏层输出上，选择同一个target（T = {t1, t2, … ,tm}）的所有hidden output：H={ht1, ht2, … ,htm}，得到target的向量表征Vt：

是target的self-attention向量，使用的模型是双层感知机。

* Sentence-level的attention建立在encoder的隐藏层输出，Target-level层和aspect向量上，句子的每一个hidden output H={h1, h2, … ,hL}都经过一个多层神经网络模型（激活函数为tanh和softmax）得到dm维的向量，此向量包含了对应词的信息，最终得到一个句子所有词的概率分布：（va是aspect的词向量，⊙ 是拼接 vt to each hi，即target的hidden output增加了attention维度，其他词不变）

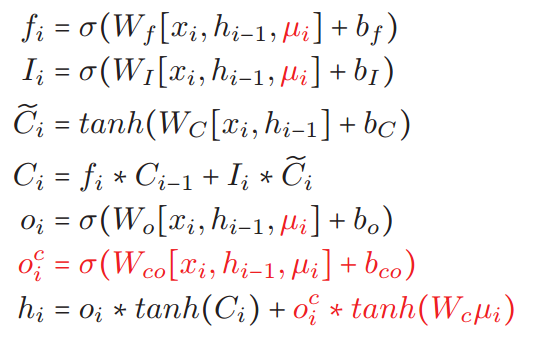
⊙

根据此概率分布和对应的hidden output句子的attention向量：（s表示句子，t表示target，a表示aspect）

每一对【target，aspect】对应一个sentence向量，即一个target有n个对应的基于aspect的句子向量，如：【west London，price】【west London，safety】，把这些句子向量输入到多类分类器（如smp的31分类），然后得到针对特定属性的标签结果（sentiment polarity）（即none，中立，正，负等），即转换为多类问题。

* 嵌入知识的LSTM
* 知识在此处充当的角色是：1）帮助信息流入下一个时间步；2）给记忆单元补充额外的背景知识。首先有个针对当前句子的候选知识集K= {μi,1, μi,2, ⋯, μi,K}，用dc维向量表示每个知识，得到表示知识的向量：

再将其拼接到LSTM的输入中（并未改造LSTM的结构）：



模型的损失函数是基于【target，aspect】对的交叉熵：



1. 疑问
2. 涉及的需要关注的模型（方法）

* Xu et al. 2016——“adds a knowledge recall gate to the cell state of LSTM.”

1. 知识拓展
2. 体会及思考

知识与主题词是同等重要还是有等级

包含一个知识和多个知识（同类别和不同类别两种情况）的不同处理：如按类别叠加向量

知识的词向量怎么处理，或者说需要不需要知道知识的词向量，还是仅需知道类别即可（只知道类别即可还属于知识的应用吗）

主题词一般有词向量，可以将【主题词-知识类别】对的信息编入（到attention）

采用增加数据信息/维度：附加的知识仅在target处加（在哪一层）；每一对【主题词，知识】得到一个句子向量，再进入小分类器；在每个门都加；【知识-类别】向量；