

利用外部知识提升 aspect-level 情感分析

摘要

Aspect-level 情感分析是用来识别一句话中某一个 aspect 对应的情感极性，其中 aspect 可以是 aspect term, 或者对应的 aspect category。如给定句子“the salmon is tasty while the waiter is very rude” 其中的 term 分别是 salmon 和 waiter，对应的 aspect category 分别是 food 和 service，对应的极性分别是 positive 和 negative。研究人员使用深度学习的方法针对 aspect-level 的研究已经取得了可观的成果，然而研究中使用的数据集较小，同时 aspect-level 数据集标注不仅困难而且成本很高，这极大的限制了深度模型的能力。为了解决这一问题，研究人员引入了外部知识进行数据量的扩充，从而提高情感分类的性能，其中包括情感相关的常识知识，document-level 知识。此外，还有一部分研究人员使用联合学习的方式将 aspect extraction 和 aspect 情感分类联合起来，以减小数据标注的困难。

背景

文本情感分析按照粒度可以分为三个级别，分别是 document-level、sentence-level、aspect-level。Document-level 情感分析是假设每个文档表达一种情感，句子级情感分析是将整个句子作为分析对象，分析其整体的情感。这两种粒度的情感分析都存在两个问题，一是不能得到更加准确和细节的信息，不知道针对具体方面的情感，二是段落或者句子针对不同的方面可能包含相反的情感。这就需要更加细粒度的情感分析，aspect-level 的情感分析致力于判断每个被提及的 aspect 的情感极性，因此能够得到更加准确和细节的信息。如上面的例子：“the salmon is tasty while the waiter is very rude”，针对 salmon 和 waiter 的情感极性分别是积极和消极的情感，然而如果以句子为粒度进行分析，则可能得到中性的情感极性，这将造成情感极性的不精确判定。Aspect-level 情感分析具有很大的实践意义，如上一例子所示，它可以用来帮助寻找哪里 salmon 比较好吃的人提供建议，从而帮助他们快速的做出决策，针对 waiter 表达出的消极情绪可以帮助餐馆的管理者提升餐馆文化的培养教育，此外，对用户评论中表达出对特定 aspect 的情感信息，可以更好的构建用户画像，从而用于推荐系统得到精准的推荐。因此，这次本文主要针对 aspect-level 情感分析展开调研。

Aspect-level 情感分析从 2014 年国际语义评估组织（international Workshop on Semantic Evaluation）发布相任务 4 之后备受关注，发布的这个任务分为四个相应的子任务，分别是 aspect term extraction、aspect category extraction、aspect term sentiment analysis、aspect category sentiment analysis。研究人员通常使用相应的数据集将这些子任务分开进行研究，但也有部分研究人员为了使研究更加具有实际意义而将 aspect term extraction 和 aspect term sentiment analysis 利用任务之间的相似性一起进行研究。

针对 aspect-level 情感分析的方法分别有基于情感词典的、使用传统监督学习的以及基于深度学习的。基于情感词典的方法需要为每个词分配情感值，并将 aspect 关联的词的情感值聚合为该 aspect 的情感值，这种方法需要预先准备情感词字典。传统的监督学习算法很大程度上依赖于人工设计的特征，这些方法不仅需要耗费大量的人力物力，而且不能捕捉 aspect 和其上下文之间的语义关系，由于深度学习的强大的表征能力及其在 NLP 一些领域的成功应用，研究人员将其

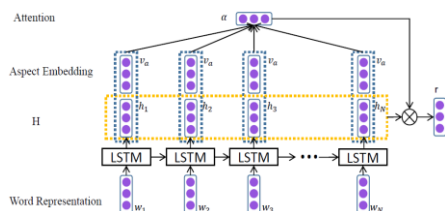
引入 aspect-level 情感分析研究中。

早期的利用深度学习对 aspect-level 情感分析的研究并没有将 aspect 信息考虑在内，这也引起了深度学习模型错误的识别句子中信息的问题。基于此观察，研究人员(Wang et al. 2016; Tang et al. 2016) 将 aspect 信息考虑在内，通过编码后将其信息融入模型中，并通过可以得到相对重要信息的注意力机制进一步查询和此 aspect 相关的上下文，以此来构建整个输入句子的表示，从而进行情感分类任务。在此基础上，研究人员(D. Ma et al. 2017)指出先前存在的方法忽视了 aspect 的独立建模，他们认为 aspect 及相应的上下文应该交互的进行学习，不仅从 aspect 信息中得到对应重要的上下文表示，也应从上下文词中寻到对应重要的 aspect 词。这些基于深度学习模型的研究方法已经取得了不错的成果，然而，通过观察他们研究所用的数据集，我们可以发现这些数据集数量很少，这极大的限制了深度模型的能力。为了解决此问题，大多数人最直观的想法便是多增加一些数据，可是 aspect-level 数据集标注困难，因为不同的人对于所表达的情感理解不同，此外，标注成本也很高。因此，一些研究人员尝试使用一些相对容易获取的外部知识，如：情感相关的常识知识(Y. Ma et al. 2018)、document-level 知识(He et al. 2018; Z. Li et al. 2019)引入 aspect-level 情感分析问题中以提升深度模型对其分类的性能。此外，一些研究人员(X. Li et al. 2019; Hu et al. 2019; D. Ma et al. 2017)通过联合学习的方式，将 aspect term extraction 和 aspect-level 情感分析任务结合用来解决数据标注困难的问题。

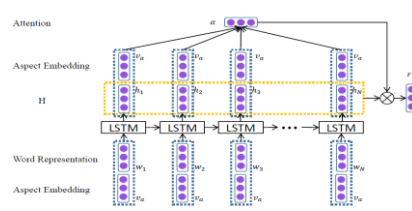
本文接下来将分别介绍使用深度学习模型进行 aspect-level 情感分析的三种经典模型、为弥补数据补足引入外部数据的方法以及通过联合学习的方式解决数据不足问题的方法。

基于深度学习模型的三种经典的方法

先前基于深度学习模型对 aspect-level 情感分析问题的研究中并没有将 aspect 信息考虑在内，这就引起了模型错误的识别不相关信息的问题。为了解决此问题，(Wang et al. 2016)指出句子中的情感极性不仅由句子中的内容决定同时由句子中涉及到的 aspect 决定，因此，作者将 aspect 信息加入模型中，同时引入了注意力机制，用于识别特定 aspect 输入时对应重要的上下文，并利用此信息生成最终的句子表示，通过 softmax 层进行情感极性的分类。为了利用 aspect 信息，作者提出了 AT-LSTM，如图一所示，将 aspect 的信息作为注意力机制的查询项，用于判定相应上下文的注意力权重，以此来构建句子的表示，此外，作者为了更加充分的利用 aspect 信息，进一步地提出了 ATAE-LSTM 模型，如图二所示，将 aspect 编码后的信息直接和句子中的每个词进行拼接之后通过 LSTM 得到对应的隐状态，再做同 AT-LSTM 相同的操作，其模型图如下：



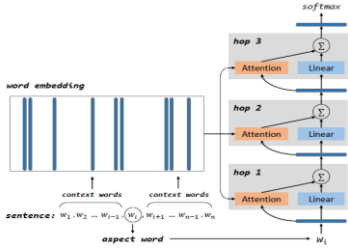
图一 AT-LSTM



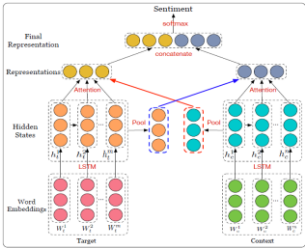
图二 ATAE-LSTM

作者通过实验验证了加入 aspect 信息及引入 attention 机制提升了情感分类的准确度。

同样地, (Tang et al. 2016) 提出 Memnet 模型架构, 如图三所示, 使用 memory 网络替代 LSTM, 将句子的上下文信息存于 memory 中, 使用编码后的 aspect 信息作为查询机制, 通过层级的 attention 机制, 构建 aspect 最终的表示, 通过 softmax 进行最终的情感分类。值得一提的是是, Memnet 中使用的注意力机制不仅是基于内容的注意力机制, 而且作者将位置信息引入进来, 这里作者认为, 距离 aspect 词越近的上下文词应该对 aspect 的影响越大, 这一点被后来的研究广泛使用。



图三 Memnet 模型



图四 IAN 模型

然而, (D. Ma et al. 2017)指出之前存在的研究仅仅考虑了使用 aspect 信息选择重要的上下文词以此来构建句子的表示, 然而他们忽视了上下文词对 aspect 的影响, 比如 ‘food seafood fish’ 中 fish 的权重就应该高一点, 因此, 作者构建了一个交互网络 IAN 模型, 如图四所示, 左右两边是相同的网络, target(aspect)信息和上下文信息分别通过两个网络的词嵌入层和 LSTM 层得到对应的隐状态, 通过平均池化得到平均隐状态, 使用这个平均隐状态向量作为注意力机制的查询项, 分别查询对应的重要信息, 得到相应的表示后将它们进行拼接作为句子的表示, 输入 softmax 中进行最终情感极性的判断。通过实验, 作者论证了在之前研究的基础上加入 context 词对 aspect 词权重选择的注意力机制后, 情感分类的准确度有提升。这种交互思想也被后来的研究广泛使用。

在这一部分, 将以上的三篇文章中提到的模型进行实验结果的对比, 结果如表一所示:

表一：三种经典模型的效果对比（Accuracy）

Models	restaurants	laptop
TD-LSTM	75.6	68.1
ATAE-LSTM	77.2	68.7
MemNet	80.95 [MemNet(9)]	72.37 [MemNet(7)]
IAN	78.6	72.1

从实验结果中可以看出, 在 semEval2014 restaurant 及 laptop 数据集上, 加入了 aspect 信息及引入注意力机制的 ATAEE-LSTM 确实比没有加入这些信息的效果有提升, 同时利用 context 词选择重要的 aspect 信息, 对情感分类的准确度也有提升, 其中, memNet 的效果最好, 因为模型不仅使用了 memeoray 网络, 层级的 attention 机制, 而且将位置信息考虑在内。

从实验结果中也可以看出, 深度模型已经取得了不错的效果, 然而根据(Zhou et al. 2019)整理的常用数据集, 如表二所示, 我们可以发现, 这些数据集很小, 除了 twitter 数据集包含 7000 多条数据外, 其他的数据集都只包含 2000 多条数据, 这极大的限制了深度模型的能力, 为了解决这个问题, 我们最直观的想法就是增加数据, 可是如果使用标注的方式, 因为不同的人对于句子所表达的情感是不同的, 因此, 这种标注不仅困难, 而且成本很高, 因此研究人员就尝试将已经存在并且很好获得的外部数据引入进来, 用于提升 aspect-level 的情感分类性能。

其中包括情感相关的常识知识，document-level 知识。此外，还有一部分研究人员使用联合学习的方式将 aspect extraction 和 aspect 情感分类联合起来，以减小数据标注的困难。

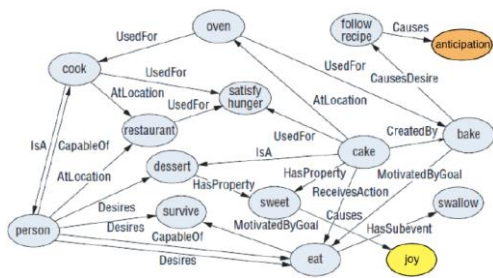
表二：常用数据集统计

引入常识知识提升 aspect-level 情感分析的方法

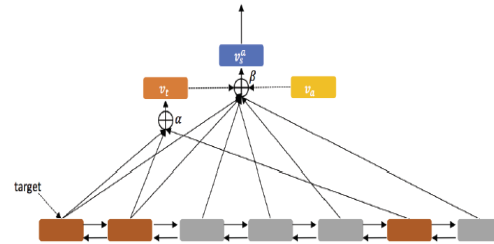
常识知识是人们认为理所当然，在生活中经常被使用，且在社会实践中往往

Dataset		#Samples	#AvgLen	#Term Set	#AvgTerm Len	#ATPS	Neg./Neu./Pos
Laptop14	train	1462	18.5855	939	1.9191	1.5821	866/633/2,164
	test	411	14.9562	389	1.9434	1.5523	128/169/341
Restaurant s14	train	1978	16.2856	1191	2.0722	1.8210	805/633/2164
	test	600	15.4167	520	1.9942	1.8667	196/196/728
Restaurant s15	train	1120	13.1009	492	2.0163	1.4768	403/53/1198
	test	582	14.3728	252	1.8968	1.4519	346/45/454
Restaurant s16	train	587	13.5427	671	2.0596	1.4678	749/101/1657
	test	587	13.4957	289	1.8581	1.4634	204/44/611
twitter	train	6248	18.8078	113	1.7965	1.0	1560/3127/1561
	test	692	18.8671	82	1.8049	1.0	173/346/173

假定人们已经掌握但往往不会在文本中进行描述的知识。人类在面对新情况时，通过已经学习的知识和经过社会实践积累的一些经验可以做出精准的决策，而机器在面对新情况时，因为缺少一些给定的规则或泛化能力因此不能做到举一反三甚至有时会造成灾难性的错误。因此，(Cambria et al., 2016)认为将关于情感的常识信息加入情感分析的研究中，效果理所应当有提升。他们通过概念元的方法构建了一个情感相关的常识库，部分缩略图展示如图五，展示了关于 cake 的情感常识部分。基于他们构建的情感相关的常识库，(Y. Ma et al., 2018)将其加入到 target-level（这篇文章的 target 即 aspect）情感分析中。作者使用 AffectiveSpace 在不损失原始空间的语义及情感关系前提下将图五所示的高维数据映射为低维的连续嵌入向量，并将其添加到 LSTM 构成 senticLSTM，用于 1）帮助时间步之间的信息流动 2）提供记忆单元的补足信息，接着作者使用了 target-level 的注意力机制，使其利用自注意力机制选择出 target 词中比较重要的词，将其和对应的 aspect（即 category）做拼接，随后进行 sentence-level 的注意力机制，以此构建整个句子的表示，并将结果输入到 softmax 层进行情感分类。



图五 关于‘cake’的情感常识部分



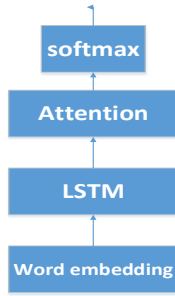
图六 SenticLSTM+SA+TA

这篇文章引入了情感相关的常识知识值得借鉴，这里作者将高维的图映射为向量的形式难免会造成部分的信息丢失，此处可以尝试使用 GNN 网络进行进一步的探讨。

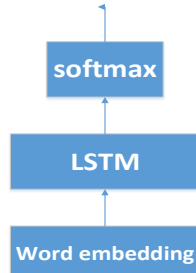
引入外部文档级别知识提升 aspect-level 情感分析的方法

因为文档级别知识相对容易获取，并且如 yelp、amazon 等文档级别的知识存在评分，因此很容易将其转换为对应的情感极性。(He et al. 2018)首次将文档级别的知识引入到 aspect-level 情感分类问题中，作者使用的便是 yelp2014 数据集及 Amazon Electronics dataset,其中，在训练过程中，根据领域特定的知识，作者将 yelp2014 用以辅助 restaurant14/15/16 数据集，将 Amazon Electronics dataset 用以辅助 Laptop 数据集进行训练，作者分别将这两个数据集对应的评分转换为情感极性，其中，评分>3 的为积极，=3 的为中性，<3 的为消极。作者提出两种迁移方法使用文档级别的外部知识进行训练：1) 预训练方法，2) 多任务方法。预训练是计算机视觉中常用的一种技术，在这种技术中，低层神经层可以有效地转移到不同的任务中。多任务是通过结合的目标函数同时训练多个样本中的数据，可以提升模型的泛化能力。在预训练方法中，作者构建了两个模型，分别用于 aspect-level 及 document-level 的情感分类，其模型示意图大致如图七、八所示，首先使用文档级别的数据在图八所示的架构中进行训练，将 LSTM、word embedding 层训练好的参数对图七所示的模型架构对应的 LSTM、word embedding 层的参数进行初始化，之后 aspect-level 情感分类模型再在自己的数据集上对每层的参数进行训练，最终得到所有的参数用于最后的 aspect-level 的情感分类任务。在多任务方法中，作者将 LSTM、word embedding 作为两个任务的共享层，而在输出时根据特定的任务进行独立的输出，模型根据整体的损失函数进行训练，通过任务共享层，模型提高了泛化能力。最后，作者进一步将两种方法进行结合，首先通过预训练的方法训练文档级别情感分类模型的参数，使用其参数对多任务模型对应的层进行参数初始化，之后通过数据进行多任务模型对应的参数微调，进而得到最终的参数用于 aspect-level 的情感分类。

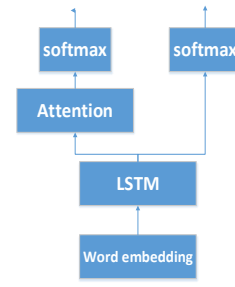
这篇文章是首次将文档级别的数据集应用到 aspect-level 情感分类任务中，并在预训练+多任务模型上取得了很好的提升。实验结果证明了随着引入外部文档比例的增加 Accuracy 和 Macro-F1 均有增加。



图七 aspect-level sentiment Classification model

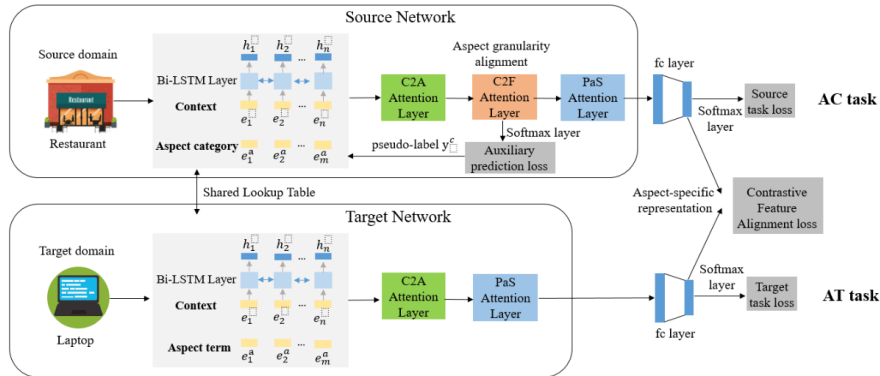


图八 document-level sentiment Classification model



图九 multi-task learning

(Z. Li et al. 2019)认为虽然 aspect-term 标注存在困难,但 aspect category 标注就相对简单,且 aspect category 一般包含在可数的集合中,如 restaurant 领域中特定的 AC 集合有{“food”, “service”, “speed”, and “price”},因此收集用户对于不同 aspect 类别的偏好是符合实际并且容易获取的。因此,作者构建了一个大规模、跨领域的 AC-level 数据集,在构建数据集时作者首先使用工业级的 opinion parser 解析器识别 aspect category 及其情感标签,然后手工的重复检查由 op 系统错误的标注项并有意识地选择更多的否定、对比和疑问句,使其更具挑战性。作者使用此数据集构建粗粒度的 aspect category 情感分类任务,此外,作者用此高资源领域粗粒度的 aspect category 情感分类任务去提升低资源目标领域的细粒度 aspect term 情感分类任务,作者提出了多粒度对齐网络 (Multi-Granularity Alignment Network) 解决粒度不一致和领域之间特征不匹配的问题,其模型架构如图十所示,MGAN 包含两个网络分别学习两个领域的 aspect 特定的表示。两个网络分别通过 word embedding 及双向 LSTM 得到对应的隐状态表示后,分别进入多个注意力机制,其中 Context2Aspect(C2A) Attention 是根据上下文确定源领域中给定 aspect category 及目标领域中给定的 aspect 每个词的注意力权重。Coarse2Fine (C2F) Attention 根据 aspect category,找出 aspect term 实现粒度对齐并提供位置信息以供下一步使用。Position-aware Sentiment (PaS) Attention 是基于和 aspect term 距离比较近的情感词更有可能是 aspect term 的修饰词的值观想法,由位置信息选择出 aspect 对应的重要的上下文信息,构成句子的最终表示。Contrastive Feature Alignment (CFA)通过语义对齐和语义分割实现领域对齐,其中语义对齐即同类(相同情感极性)不同领域的分布尽可能相同,语义分割即不同类(不同情感极性)不同领域的分布尽可能不同。



图十 MGAN 网络框架

其中, Coarse2Fine (C2F) Attention 中,作者基于自编码器的思想,加入了一个 pseudo 标签预测的辅助任务,在做这个任务中,源 aspect category 不仅代表 aspect

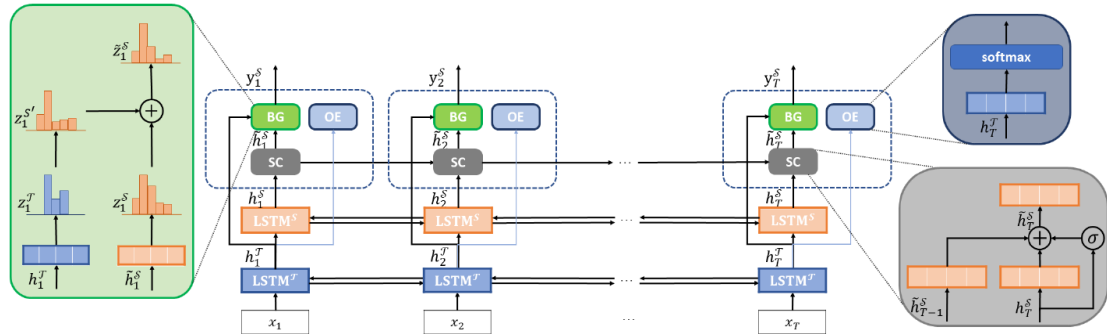
词的类别信息，而且被认为是真实的伪标签,利用上一步获得的关于上下文的注意力表示的 **aspect category** 向量和上下文向量结合反过来预测伪标签信息，如果上下文中包含和源 **aspect category** 很相关的 **aspect term**，那么为了得到好的标签预测结果，注意力机制就会强调这个 **aspect term**，作者就认为注意力分数高的即为 **aspect term**，这也为下一步提供了位置信息。如果上下文是通过隐式的表达关于这个 **aspect category** 的情感，则句子中可能不存在 **aspect term**，为了克服这一情况，作者提出了一种类似 RNN 门机制的融合门 F，自适应的控制通过 C2A 得到的表示和通过 C2F 得到的表示的比例，最终得到 **aspect term** 的表示。

通过跨领域的数据迁移，**aspect-level** 情感分类的性能得到了进一步的提升。

通过联合学习的方式提升 **aspect-level** 情感分析的方法

先前存在的方法大都是在 **aspect** 已经给定的情况下进行的研究，而在实际情况中，用户产生的一些评论并不包含这些标注信息，为了更加符合实际情况并且减小数据标注的困难，(He et al. 2019; X. Li et al. 2019; Hu et al. 2019)提出通过 **aspect term** 抽取和 **aspect term** 情感分类任务联合学习的方式提升 **aspect term** 情感分类的性能。

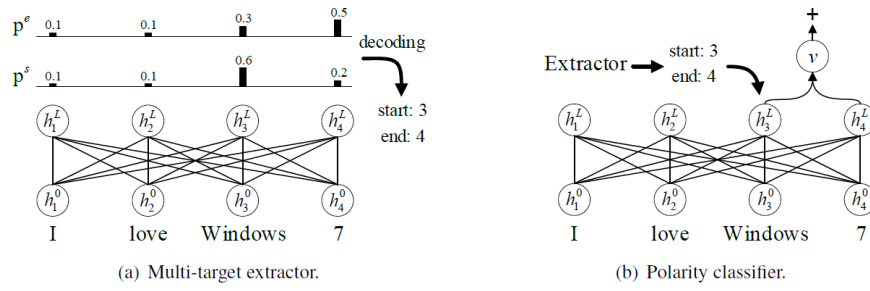
(X. Li et al. 2019)将 **aspect term** 抽取及情感极性分类任务构建成一个统一的模型，并通过统一标签的形式解决此任务。具体地，作者提出了如图十二所示的模型框架，框架包含两层堆叠的递归神经网络，下面的一层执行 **target** 的边界预测用于指导上层的网络提升情感分类的性能，上面的一层预测统一的标签生成最终的输出结果。



图十二 统一模型框架

此外，作者使用了三个辅助的组件，其中，作者基于观点目标和观点词总是同时出现，接近观点词的词很有可能使目标词的假设引入了 OE 组件，OE 组件使用观点词进一步探测 **target** 词，以精炼边界信息的隐状态表示。作者构建的 SC 组件使用一个简单的门机制使相同 **target** 包含的此保持情感一致。门机制显式的增强当前词和先前词的特征。作者使用 BG 组件构建转移矩阵编码边界概率到统一标签空间的概率分布。在做统一标签决策时，作者将一定比例的边界概率分数考虑在内并使用上层 RNN 的概率分数进行最后的情感极性的预测。

(Hu et al. 2019)认为如果采用序列标注的方式会造成搜索空间大和情感不一致问题。为了解决这一问题，文章提出了 **span-based extract then classification** 框架，如图十三所示，在目标片段边界的监督下直接从句子中抽取多个观点目标，用目标的 **span** 表示进行情感极性的分类。

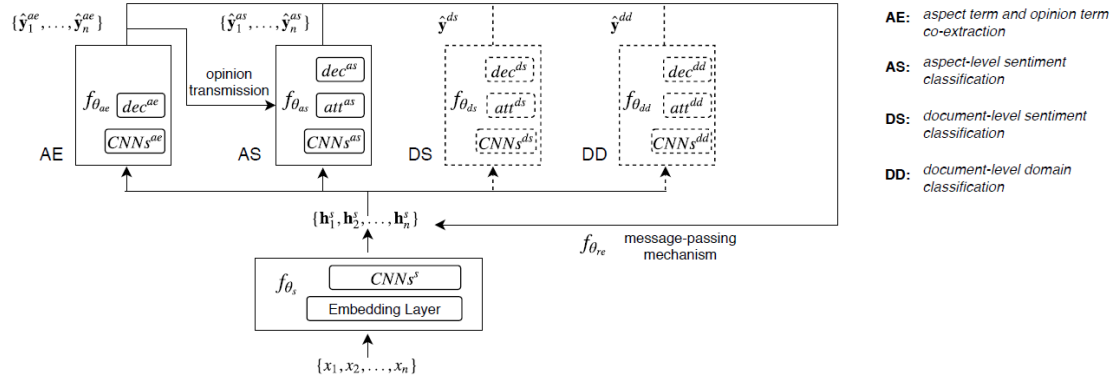


图十三 基于 span 的抽取

作者将句子通过 bert 预训练后得到对应的隐状态表示，通过构造的目标的开始和结束位置的概率分布获得对应的分数，并依据词分数得到最终的位置信息，这里作者借鉴了 CV 中非极大值抑制的思想，将抽取的重叠的 span 进行删除，以获得最佳的 span，得到位置信息后，将该位置对应的 aspect 通过基于内容和基于位置的注意力向量表示进行结合得到最终的 aspect 表示，并将其送入 softmax 中进行情感分类。

通过 Bert 的预训练方式及基于 span 的 aspect term extraction 和对应的情感分析联合学习的方式，相比于简单的序列标注方式效果有提升。

(He et al. 2019)不仅将 aspect term 抽取和 aspect term 情感极性分类任务进行联合，而且通过引入外部文档级别的知识将 token-level 和 document-level 任务进行联合，使模型同时学习多个相关的任务，作者提出的多任务学习框架 (interactive multi-task learning network IMN) 如图十一所示，和传统的依赖于学习不同任务之间的共同特征的多任务学习方式不同，这里作者提出了信息通过一个共享的隐变量集合迭代的在不同的任务之间进行训练的信息传递框架。



图十一：IMN 框架

在作者提出的模型中，最下面一层是所有任务共享的特征抽取组件，特征抽取组件将句子通过 embedding 层及 CNN 层转换为对应的隐状态之后分别输入 aspect term 抽取、aspect term 情感极性分类、文档情感极性分类、文档领域分类四个任务中。在 AE 任务中，共享的隐状态通过 CNN 编码后得到任务特定的隐状态，和之前共享的隐状态进行拼接后输入 softmax 进行分类，其输出的 label 为输入的对象是否属于 aspect 或者 opinion 项，使用五类标签 {BA, IA, BP, IP, O} 进行标注，五类标签分别代表 aspect 的开始和内部、观点项的开始和内部、其他的词。

在 AS 任务中，通过 CNN 层得到的隐向量结合 AE 任务的输出后一起输入到自注意力层，以使 AE 任务中预测的 opinion term 辅助 AS 任务的情感分类任务。其中，在计算自注意力权重时不仅使用到了 token 是 opinion term 的概率信息，同时使用到了 token 之间的距离信息，随着 token 距离的增加，相互之间的

影响力降低。将通过自注意力层得到的表示和最初的共享表示做拼接后作为输入 token 的最终表示，通过 softmax 之后输出代表其情感极性的标签信息，使用 {pos,neg,neu} 三类标签进行标注，分别代表积极、消极、中性的情感极性。

对于文档级别的情感分类任务 DS 任务，其目的是对于输入的文档进行情感极性的预测。对于文档级别的领域分类任务，其目的是对输入的文档进行领域标签的预测。两个任务中均包含 CNN 层、注意力层及最终的分类预测层。

其中，作者模型中构建的信息流动机制是将上一个时间步的共享表示和 AE AS DS 任务的输出及 DS DD 中的注意力权重进行拼接，其中 DS DD 中的注意力权重反应第 i 个 token 的情感相关及领域相关信息，作者认为 token 情感相关及领域相关性越强，则极有可能是 opinion 词或者 aspect 词，并认为这些信息可以辅助 aspect level 的任务。

通过 aspect term 抽取及对应的情感分类任务联合训练，加上引入外部文档知识的辅助，模型取得了很好的效果。

总结

表三：实验结果对比及分析

Models	Rest14		Rest15		Rest16		Laptop14		Twitter	
	ACC	F1	ACC	F1	ACC	F1	ACC	F1	ACC	F1
TD-LSTM	75.6						68.1			
ATAE-LSTM	77.2						68.7			
MemNet	80.95 [9]						72.37[7]			
IAN	78.6						72.1			
+commonsense			76.47							
+document-level +Pretraining +multi-task learning	79.11	69.73	81.30	68.74	85.58	69.76	71.15	67.46		
MGAN	81.49	71.48					76.21	71.42	74.62	73.53
Unified Model		14/15/16 F1: 69.80						57.90		48.01
Span-based model		14/15/16 F1: 74.92						68.06		57.69
IMN	83.89	75.66	85.64	71.76			75.36	72.02		

从表三整体的实验结果中，我们可以看到，引入外部知识确实对 aspect-level 情感分类性能有提升。具体的，引入大量文档级别对应领域的知识比引入情感常识相关的知识提升明显，一方面，可能是因为引入常识相关的论文中，作者并没有使用最新提出的情感相关的常识库，另一方面，大量领域相关的文档级别的知识不仅解决了 aspect-level 数据不足的问题，而且解决了数据中情感极性不均衡的问题。MGAN 模型不仅引入了文档级别的外部知识，同时，作者实现了跨领域之间的数据辅助训练功能，作者通过对 aspect category 数据进行标注实现的跨粒度及跨领域的文档知识迁移在直接引入外部数据方法中取得了最好的效果。

在将 aspect term 抽取和对应的情感极性分类任务联合学习训练解决数据标

注困难问题的几种方法中，使用统一模型的方法在 restaurant 数据集上表现的比直接引入外部数据效果要好，原因可能是通过联合学习的方式，模型学习到了未进行标注但实际存在的一些 aspect term。基于机器翻译思想，通过 span-based term 抽取的方式解决了序列标注进行联合学习中的搜索空间巨大和情感不一致问题，进一步的提升了联合学习的 aspect-level 情感分类性能，除了使用 span-based term 抽取方法，作者构建的模型使用了 Bert 的预训练方法，这也可能是性能提升的关键一步。IMN 模型不仅构建了 aspect term 抽取及对应的情感分类联合训练的方法，同时作者将文档级别的知识进行迁移，通过构建 aspect term extraction、aspect term sentiment classification、document-level sentiment classification、document domain classification 四个任务一起学习的方法使 aspect-level 情感分类性能达到了最好的效果。

整体上，引入 aspect 信息及注意力机制相比于之前的研究方法更准确的关注了上下文信息，因此效果有所提升。直接引入外部知识的方法，增加了数据量同时解决了数据不均衡的问题，进一步将效果进行提升，通过 aspect term 抽取及对应的情感极性分类联合学习的方式可以抽取出未进行标注但实际存在的 aspect term 从而提升了实验效果。

展望

前面介绍的几种方法已经论证了通过外部知识的引入，可以使 aspect-level 情感分类性能提升。通过先前的一些工作，我们可以发现，1) aspect-level 情感分类任务借鉴了阅读理解任务中的某些思想，那我们是否可以将其转化为阅读理解任务，使用阅读理解任务中大量的有标注数据，利用具有统一表示的有标注数据进行 multi-task，然后再在目标任务上进行 fine-tune，效果理应会有提升。2) 因为情感相关的常识知识已经被证明在引入 aspect-level 情感分类任务后会对其效果有提升，之前的工作是在其他人已经构架好的常识相关的数据集上进行工作，那我们是否可以利用其他任务上已经存在的数据集进行性能的提升？如常识推理中的心理推理方面的数据集，Triangle-COPA、Story Commonsense、Event2Mind 等，这些数据集均是关于情绪（情感）和意图的推理。此外，3) 通过观察我们可以发现，现在人们在聊天、写评论过程中经常会使用大量的表情包、emoji 等，对涉及到的这些表情包、emoji 等进一步探索，使其结合文本对潜在的 aspect 进行探索及对 aspect 的情感极性进行分析，也是下一步可以深入研究的方向。

参考文献

- Cambria, Erik, et al. "SenticNet 4: A semantic resource for sentiment analysis based on conceptual primitives." Proceedings of COLING 2016, the 26th international conference on computational linguistics: Technical papers. 2016.
- He, Ruidan, Wee Sun Lee, Hwee Tou Ng, and Daniel Dahlmeier. 2018. "Exploiting Document Knowledge for Aspect-Level Sentiment Classification." ArXiv:1806.04346 [Cs], June. <http://arxiv.org/abs/1806.04346>.
- He, Ruidan, Wee Sun Lee, Hwee Tou Ng, and Daniel Dahlmeier. 2019. "An Interactive Multi-Task Learning Network for End-to-End Aspect-Based Sentiment Analysis." In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 504–15. Florence, Italy: Association for Computational Linguistics. <https://doi.org/10.18653/v1/P19-1048>.

- Hu, Minghao, Yuxing Peng, Zhen Huang, Dongsheng Li, and Yiwei Lv. 2019. "Open-Domain Targeted Sentiment Analysis via Span-Based Extraction and Classification." In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 537–46. Florence, Italy: Association for Computational Linguistics. <https://doi.org/10.18653/v1/P19-1051>.
- Li, Xin, Lidong Bing, Piji Li, and Wai Lam. 2019. "A Unified Model for Opinion Target Extraction and Target Sentiment Prediction." Proceedings of the AAAI Conference on Artificial Intelligence 33 (July): 6714–21. <https://doi.org/10.1609/aaai.v33i01.33016714>.
- Li, Zheng, Ying Wei, Yu Zhang, Xiang Zhang, and Xin Li. 2019. "Exploiting Coarse-to-Fine Task Transfer for Aspect-Level Sentiment Classification." Proceedings of the AAAI Conference on Artificial Intelligence 33 (July): 4253–60. <https://doi.org/10.1609/aaai.v33i01.33014253>.
- Ma, Dehong, Sujian Li, Xiaodong Zhang, and Houfeng Wang. 2017. "Interactive Attention Networks for Aspect-Level Sentiment Classification." ArXiv:1709.00893 [Cs], September. <http://arxiv.org/abs/1709.00893>.
- Ma, Yukun, Haiyun Peng, and Erik Cambria. "Targeted aspect-based sentiment analysis via embedding commonsense knowledge into an attentive LSTM." Thirty-Second AAAI Conference on Artificial Intelligence. 2018.
- Tang, Duyu, Bing Qin, and Ting Liu. 2016. "Aspect Level Sentiment Classification with Deep Memory Network." In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, 214–24. Austin, Texas: Association for Computational Linguistics. <https://doi.org/10.18653/v1/D16-1021>.
- Wang, Yequan, Minlie Huang, xiaoyan zhu, and Li Zhao. 2016. "Attention-Based LSTM for Aspect-Level Sentiment Classification." In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, 606–15. Austin, Texas: Association for Computational Linguistics. <https://doi.org/10.18653/v1/D16-1058>.
- Zhou, Jie, Jimmy Xiangji Huang, Qin Chen, Qinmin Vivian Hu, Tingting Wang, and Liang He. 2019. "Deep Learning for Aspect-Level Sentiment Classification: Survey, Vision, and Challenges." IEEE Access 7: 78454–83. <https://doi.org/10.1109/ACCESS.2019.2920075>.