|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 论文题目 | Modelling Context and Syntactical Features for  Aspect-based Sentiment Analysis | | | |
| Paper URL | https://www.aminer.cn/pub/5ec49a639fced0a24b4de86a/modelling-context-and-syntactical-features-for-aspect-based-sentiment-analysis | | | |
| Project URL |  | | | |
| 综述/背景介绍 | 以前不足 | 现有改进 | 意义 | 关键词（速记词汇、信息索引词汇） |
| 以前的ABSA的不足（基于单词嵌入模型）：  1、没有充分利用句法信息的重要性。因此，方面提取模型往往不能检测到多词方面术语的边界。  2、方面情感分类器无法解释方面词与上下文词之间的句法相关性 。  3、大多数研究都是基于方面提取和方面情感分类中的一个方面。  现在研究（基于上下文嵌入的模型）：  1、从句子的语法方面，并运用自注意机制进行句法学习。我们结合了词性嵌入、基于依赖的嵌入和情境化嵌入(例如BERT、RoBERTa)来增强方面提取器的性能。  2、我们还提出了句法相对距离，以减轻不相关词的不利影响，这些词与方面词的句法联系较弱。这提高了方面情感分类器的准确性。  意义：结合方面提取和情感分类提高对句子的情感分析的准确性。 | | | 方面提取、情感分析、情感分类、自然语言处理 |
| 假设 |  | | |  |
| 方法描述(含图) | 5  方法模型    ABSA框架图  相关准备：  1、词性嵌入模型：词（数据处理层）+特征学习（后续层）  2、情境化预训练语言模型：  1）、BERT 是一种屏蔽语言模型(LM)。  2）、RoBERTa是对BERT模型进行了改进。  3）、AEN-BERT 使用BERT嵌入上下文序列和目标序列。  4）、LCF（局部上下文焦点）BERT 采用上下文动态掩蔽/上下文动态加权方法，利用语义相对距离对情感信号进行定位。距离是通过上下文词和目标方面词之间的字数来衡量的  第一大类：方面提取AE  6  CSAE总体架构  1、输入表示：开头使用特殊的分类令牌[CLS]-结尾使用分隔符[SEP]---形成结构化形式：“[CLS]”+输入序列+“[SEP]”  2、词性嵌入：一个词法分析的工具POS来得到句子对应的词性标签序列。https://universaldependencies.org/u/pos/  3、基于依赖关系的嵌入：不是使用线性词袋上下文来形成上下文窗口，而是根据单词参与的句法关系使用基于依赖的上下文。利用依赖解析式来解析句子。  第二大类：情感分类ASC  9  拟定的LCFASC整体框架  1、输入表示:全局语境G = [CLS]+S+[SEP]+A+[SEP]。  局部语境L = [CLS] + S + [SEP]  2、CDM:即为context dynamic mask，是指根据与属性词之间的距离来屏蔽句子中的一些单词，而完全屏蔽某些单词可能过于绝对。  3、CDW:（context dynamic weighting）是按照与属性词之间的距离为单词赋予一个0到1之间的权重。  4、SRD:单词之间的SRD是通过依赖解析树中对应节点之间的最短距离来度量的。 | | |  |
| 实验设计 | 1、数据集：semeval-14的ABSA数据集  2、基线模型：为了验证CSAE和LCFS-ASC的有限性  1）第一组模型采用流水线方法，独立训练单任务模型，并将AE和ASC的输出流水线化，构建端到端的ABSA解决方案。  对于AE选择了两个基于词嵌入的模型和一个基于上下文嵌入的模型对于ACS选择了两个基于词嵌入的模型和四个基于上下文嵌入的模型。  2）第二组由集成方法组成，旨在通过统一的标记方案同时提取方面项和确定极性 | | |  |
| 数据处理 | 输入 | 处理方式 | 输出结果 | 关键词（速记词汇、信息索引词汇） |
| 上下文句子 | CSAE:   1. 句法分析 2. 依赖关系分析 3. 预训练模型分析   LCF-ASC:   1. 本地的 2. 全球的 | CSAE:  提取的方面  LCF-ASC:  表达的情感 |  |
| 结论 | 1. 作者仅在semeval-14的ABSA数据集上进行了实验。如下表所示，使用了词法信息和句法信息后，f1值有了较大的提升（CSAE>RoBERTa-AE），句法信息带来的提升更加明显（RoBERTa-Dep>RoBERTa-POS）   12  2.对于ASC任务，如下表所示，使用了LCFS的策略之后，模型性能有了很大的提高，并且CDW优于CDM；LCFS稍微优于LCF。  12  **总结和未来：**  我们提出了一个端到端的ABSA解决方案，该方案将方面提取器和方面情感分类器流水线化。结果表明，语境化模型对句子句法结构的利用使其在自动答题和自动答题任务中都有了很大的改进。我们提出的方面情感分类器优于训练后的ASC模型，并能够创建独立于领域的解决方案。所提出的SRD允许方面情感分类器将重点放在关键情感词上，这些关键情感词通过基于依赖的结构修改目标方面术语。这些实质性的改进突出了最近上下文化嵌入模型在“理解”语法特征方面的不足，并指出了未来开发更多语法学习上下文化嵌入的方向。  我们可以尝试通过应用统一的标签方案来适应我们提出的CSAE体系结构的集成方法;从而可以同时实现面向抽取和情感分类。 | | | |

1.论文总结：

（论文名中文）：基于方面的情感分析的建模上下文和语法特征

（论文名英文）：Modelling Context and Syntactical Features for

Aspect-based Sentiment Analysis

literature Review

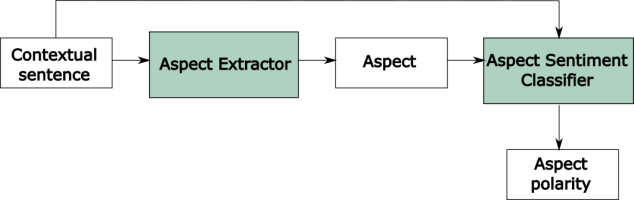
（论文URL）：https://www.aminer.cn/pub/5ec49a639fced0a24b4de86a/modelling-context-and-syntactical-features-for-aspect-based-sentiment-analysis

1. （总结）：该文是一篇技术性论文，作者没有使用一个端到端的模型，而是对属性词抽取（AE）和属性级情感分类（ASC）进行独立的研究。在AE任务中，作者结合了roberta、词法信息和句法信息；在ASC任务中，作者研究了如何在预训练语言中更好地结合句子和属性词。通过类比三种不同模型结合方法（单任务-集成任务-文章提出的方法）比较最后在semeval-14的ABSA数据集上的f1值和acc值，确定哪个方法好。结论：1.使用了词法信息和句法信息后，f1值有了较大的提升（CSAE>RoBERTa-AE）句法信息带来的提升更加明显（RoBERTa-Dep>RoBERTa-POS）

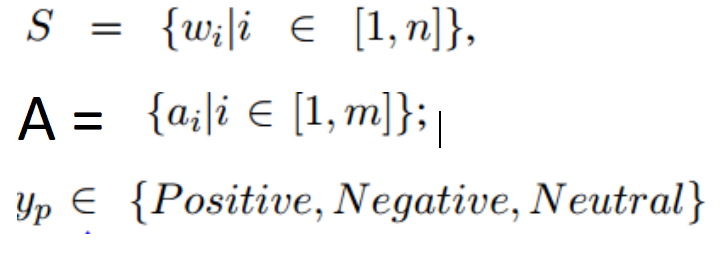
2.对于ASC任务，如下表所示，使用了LCFS的策略之后，模型性能有了很大的提高，并且CDW优于CDM；LCFS稍微优于LCF。

**重要图表：**

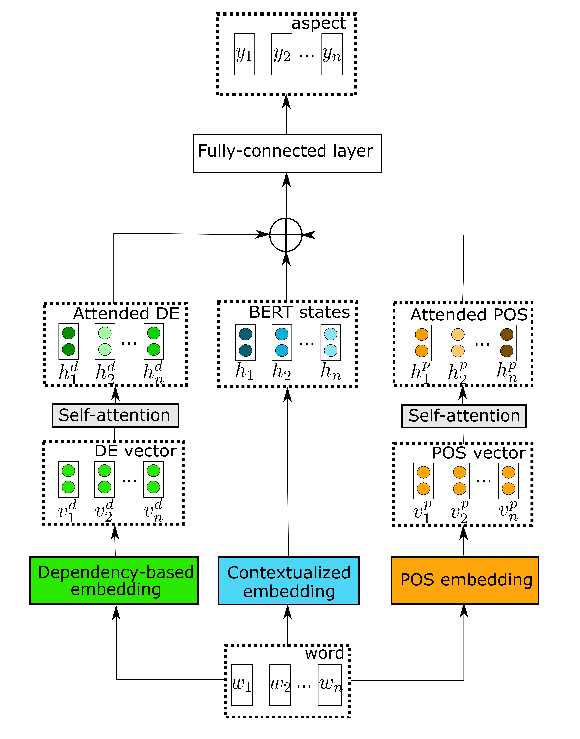
1. **研究方法：**



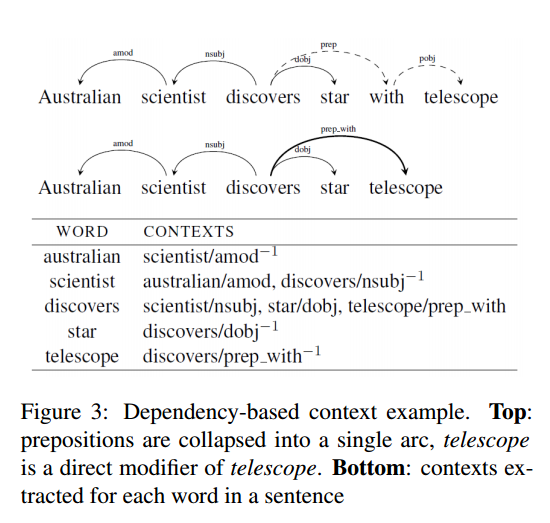
图一、ABSA框架图



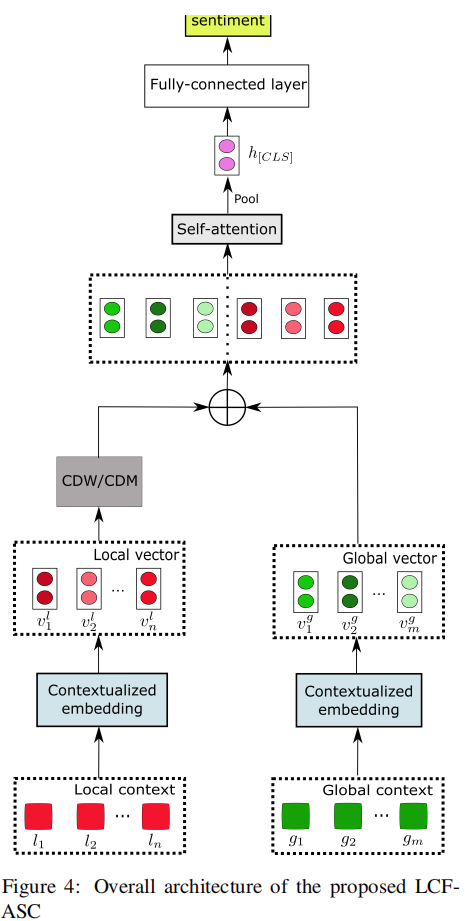
图二、方法模型



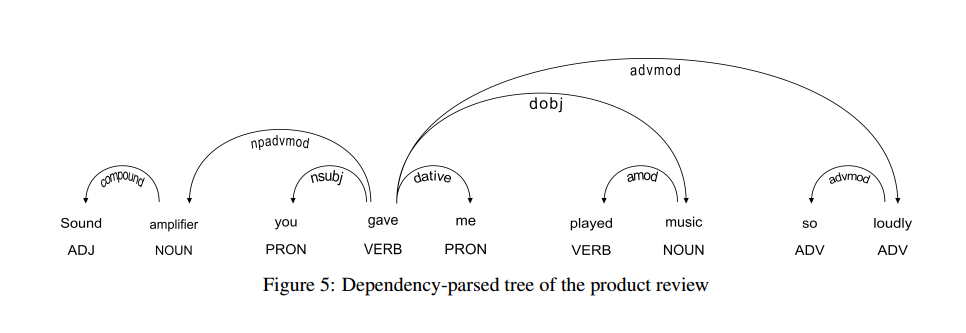
图三、CSAE总体架构



图四、依赖解析树

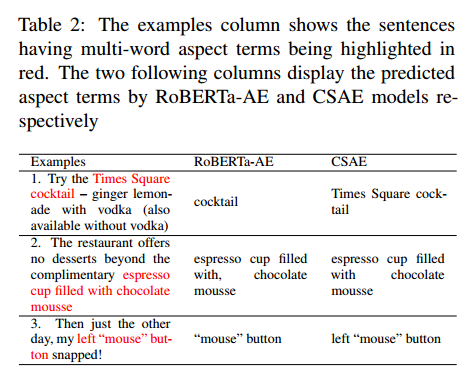


图五、拟定的LCFASC整体框架

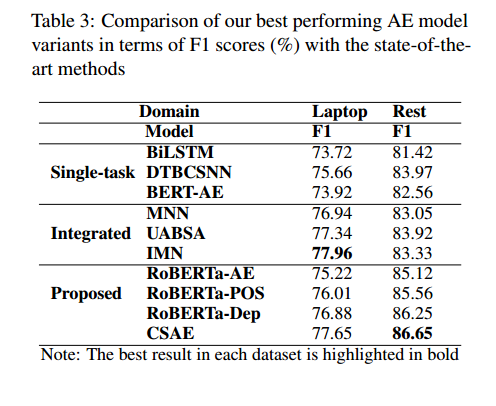


图六、依赖解析树-判相对距离

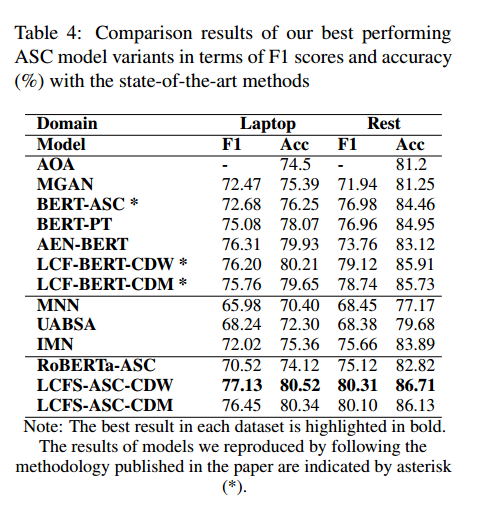
1. **研究结果：**



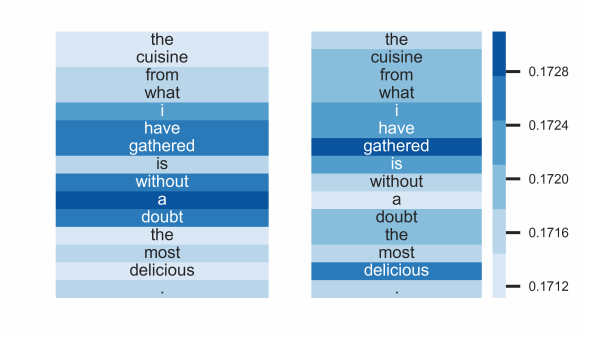
表一、RoBERTa-AE与CSAE的比较



表二、对于方面提取三方法之间的比较



表三、对于ASC三方法之间的比较



图七、视觉化注意分数对SRD影响的分析