

Effective Inter-Clause Modeling for End-to-End Emotion-Cause Pair Extraction.

*In Proc. of ACL 2020: The 58th Annual Meeting of the Association for
Computational Linguistics, pages 3171--3181*

2020ACL 的一篇论文，来自中科院自动化所团队

论文地址: <https://www.aclweb.org/anthology/2020.acl-main.289/>

代码地址: <https://github.com/Determined22>

目录

核心任务	2
与情感-原因抽取领域现有研究成果对比	2
工作主要贡献	2
模型功能概述	2
模型工作流程	3
优化部分&损失函数	3
RANKCP 模型在对候选子句对排序后面临的阈值选取问题	3
对比实验 &Baseline	4
消融实验及结论	4
结论	4
未来工作	5

核心任务

ECPE (Emotion-cause pair extraction) 任务: 直接提取文档中所有潜在的情感-原因对。

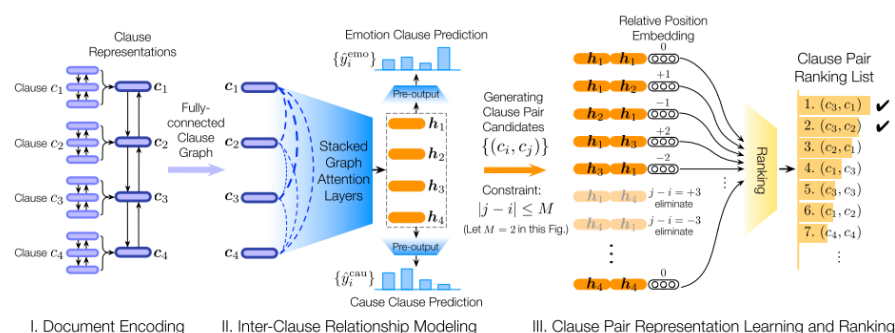
与情感-原因抽取领域现有研究成果对比

对比南京科技大学夏睿教授以及他的学生在 2019ACL 的一篇论文《Emotion-Cause Pair Extraction: A New Task》中的两步流水线方法, 本文提出的 **RANKCP** 模型是一个**端到端的联合神经网络模型**, 能够直接提取情感-原因对, 解决了两步流水线方法的错误传递问题和不能直接抽取情感-原因对的问题。

工作主要贡献

1. 提出了第一个端到端的情感原因配对提取方法, 一个从**排名 (Ranking)** 的角度来解决这一任务的统一模型;
2. 通过集成子句间关系建模和基于内核的相对位置子句对排名增强来强调子句间建模;
3. 实验结果表明, 该方法明显优于当前性能最佳的系统, 尤其是在一个文档中提取多个子句对数据的情况下。

模型功能概述



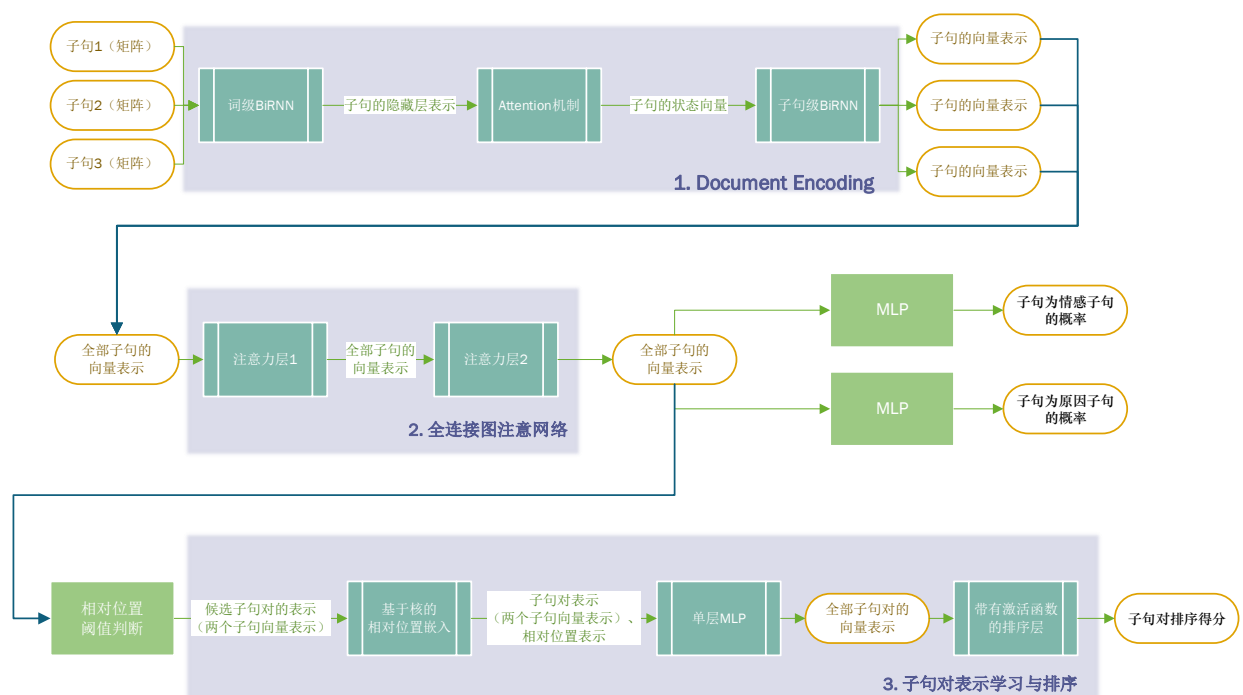
RANKCP 模型有三个模块。

第一个模块进行词级别的建模和句子级别的建模, **得到子句的向量表示**;

第二个模块构建子句间的完全图 (包括自己连接自己), 使用图注意力网络对子句进一步建模, **得到更好的子句表示**, 并且**完成两个子任务——情感预测, 原因预测**。然后两个相隔较远的句子构成情感-原因对的概率极小, 所以距离大于阈值 M 的子句对直接舍弃;

第三个模块对情感-原因对的表示加入**相对位置信息**, 得到**对的最终表示**: (子句 1, 子句 2, 距离向量)三者的拼接, 然后**进行排名**, 最终得到**排名得分**。

模型工作流程



优化部分&损失函数

RANKCP 是端到端优化的。

损失函数由两部分组成：衡量子句对的排序得分 + 衡量情感子句预测和原因子句预测结果。

$$\mathcal{L} = \mathcal{L}_{\text{pair}} + (\mathcal{L}_{\text{emo}} + \mathcal{L}_{\text{cau}})$$

这形成了对子句表示学习和子句对排名的两级监督。

RANKCP 模型在对候选子句对排序后面面临的阈值选取问题

问题：

完成排序后，很难确定一个总体阈值，以对所有文档的候选子句对进行分类（情感-原因、非情感-原因）。因为如果我们设置一个分数阈值，高于这个分数的就是情感-原因子句对，低于的不是。这种方法显然不具有普适性，有的文档可能所有对得分都不高，有的文档可能所有对得分都很高。

解决方法：使用情感极性词典(lexicon)来辅助抽取

在 Top-N 子句对中，Top-1 默认为情感子句对，其他候选子句对如果情感子句位置的子句包含了情感极性词典中的词，则将其认定为情感-原因对。因此，RANKCP 模型可以从一个给定文档中提取多个情感-原因对。

对比实验 & Baseline

1. Baseline

主要与 2019 年夏睿教授团队的论文《Emotion-Cause Pair Extraction: A New Task》中的三种方法进行对比：INDEP、INTER-CE、INTER-EC。

2. 实验的三个任务：情感-原因对抽取，情感子句抽取，原因子句抽取

RANKCP 在所有三个任务上均比其他三个 baseline 具有明显优势，其优势主要来自召回率 R 的显著改善。与 INTER-EC 相比，RANKCP 在情感原因对提取和原因子句提取方面分别实现了 8.43% 和 6.60% 的改进。这表明 RANKCP 可以有效地抽取更正确的情感-原因子句对，而不会影响精度 P。

3. 抽取多个情感-原因对实验结果

相较 INTER-EC，对于具有一对以上情感原因的文档，RANKCP 方法相对更有效。

4. ECE 任务实验结果

对于有情感标注的 ECE 任务，模型受益于子句间建模，且在原因子句抽取上表现了其有效性，可以和 CONVMS-MEMNET 模型相提并论。

消融实验及结论

进行消融实验(ablation studies)，以证明 RANKCP 的每个模块，均对提升 F1 值有帮助。

1. 采用两级监督进行训练可以提高提取性能

加入子任务（emo, cau）的监督有助于更好地学习子句表示，并最终促进子句对表示的学习和排名过程。

2. 用图注意力网络来建模子句间的潜在关系是模型的关键部分

图注意力网络是有用的；两个图注意力层效果最好，即两层能最好的建模子句间的关系。

3. 子句对表示学习对模型性能具有积极作用

4. 考虑相对位置之间的相互影响有助于获得更强大的子句对表示

去除相对位置嵌入会导致性能下降，表明子句对之间的相对位置确实对预测有用，没有相对位置嵌入的模型不能提供理想的排名列表。

结论

本文提出了第一个单步神经网络方法 RANKCP，以解决情感-原因对抽取的问题，该方法从排名的角度强调了子句间建模。RANKCP 方法有效地建模了子句间的关系，以学习子句的表示，并将相对位置增强子句对排名集成到一个统一的神经网络中，以端对端的方式提取情感-原因子句对。

在基准数据集上的实验结果表明，RANKCP 的性能明显优于以前的系统，并且进一步的分析验证了模型中每个部分的有效性。

未来工作

1. 现在对情感原因的分析主要是 clause-level 的抽取，是粗粒度，设计出细粒度的 span-level 或者 phrase-level 是不错的方向；
2. 设计有效的模型来把合适的语言知识注入到神经模型中，对情感分析任务是很有价值的；
3. 研究情感的语义角色会很有趣，这需要考虑到情感表达的全面结构并且更具有挑战性。