

面向中文文本的情绪持有者抽取研究

孙亚伟¹, 石宇¹, 韩旭^{2,*}

¹北京邮电大学, 计算器学院 (国家示范性软件学院), 北京, 100876

²中国科学技术信息研究所, 北京, 100038

{sunyawei, yus}@bupt.edu.cn, hanx@istic.ac.cn

摘要

情绪持有者是文本中带有情绪的主体, 对这些情绪持有者的分析对文本情绪理解至关重要。然而, 现有研究未充分考虑情绪持有者的共指现象, 且由于缺乏面向中文语料的情绪持有者抽取数据, 这一研究的发展受到了进一步的限制。本文构建了一个针对中文文本的情绪持有者抽取数据集, 有效解决了数据中的共指问题。同时, 提出了一种融合语义、情绪和词性特征的模型, 实现了高效的情绪持有者抽取与共指消解, 且在各项性能指标上超越了基线模型。消融实验进一步证明了模型设计的有效性。¹

关键词: 情绪分析; 信息抽取; 情绪持有者

Research on Emotion Holder Extraction for Chinese Texts

Yawei Sun¹, Yu Shi¹, Xu Han^{2,*}

¹School of Computer Science (National Pilot Software Engineering School),
Beijing University of Posts and Telecommunications, Beijing 100876

²Institute of Scientific and Technical Information of China, Beijing 100038
{sunyawei, yus}@bupt.edu.cn, hanx@istic.ac.cn

Abstract

Emotion holders are the entities in text that exhibit emotions, and their analysis is crucial for understanding emotional expressions in text. However, existing studies have not fully considered the phenomenon of co-reference among emotion holders, and the lack of Chinese corpora has also hindered the development of this research area. This paper constructs a dataset for the extraction of emotion holders from Chinese texts, effectively resolving the co-reference issues within the data. Additionally, a model that integrates semantic, emotional, and part-of-speech features is proposed. It achieves efficient extraction and resolution of emotion holders, surpassing baseline models in various performance metrics. Ablation studies further validate the rationality of the model design.

Keywords: Emotion Analysis, Information Extraction, Emotion Holder

1 引言

*通讯作者

¹本研究的数据集及代码可通过如下链接获取: <https://github.com/yaweisun/ehe>

©2024 中国计算语言学大会

根据《Creative Commons Attribution 4.0 International License》许可出版

随着互联网的迅猛发展，每天都在产生大量的文本数据。如何从这些数据中提取对人类有益的信息，成为自然语言处理（Natural Language Processing, NLP）领域的一个核心挑战。情绪在研究人类行为和心理健康方面占据了至关重要的地位(Dolan, 2002)。情绪不仅影响个体的心理状态，也会影响其决策过程、人际关系以及身体健康。鉴于文本中包含了丰富的情绪信息，因此在NLP领域，对文本进行情绪分析是一个重要的研究方向。

目前，文本情绪分析主要集中在情绪的分类和识别上，例如识别对话中的情绪(Zhang et al., 2023)以及分析社会热点事件中的情绪表现(Ding et al., 2020)。然而，这些工作仅触及情绪分析的初级阶段。为了更全面地理解文本中蕴含的情绪信息，研究者还进一步挖掘包括情绪线索(Liew et al., 2016)、情绪持有者(Campagnano et al., 2022)、情绪原因(Gui et al., 2016)以及情绪所引发的结果(Sun et al., 2023)在内的复杂情绪信息。图 1展示了这些高级情绪分析任务的实施示例。

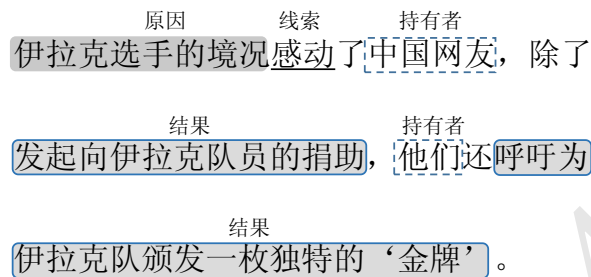


Figure 1: 文本情绪信息抽取示例

在情绪分析中，情绪线索（Emotion Cue）是用于判断文本情绪类别的关键依据。例如，在图 1中，“感动”作为情绪线索，体现了一种积极的情绪。情绪持有者（Emotion Holder）是产生情绪的主体，图中的“中国网友”和“他们”便是示例中的情绪持有者。情绪原因（Emotion Cause）是情绪的触发因素，而情绪所引发的结果通常表现为情绪持有者受情绪影响所采取的行为，即情绪驱动行为（Emotion-Driven Behavior）。这些行为是研究情绪对持有者所造成影响的重要基础，而情绪持有者则是情绪与情绪驱动行为之间的关键连接点。因此，在情绪分析中，对情绪持有者的抽取显得尤为重要。

尽管已经有一系列针对情绪持有者的研究，但这些研究通常未能充分考虑文本中广泛存在的情绪持有者表达方式的共指问题。例如，在图 1所示的示例中，“中国网友”和“他们”这两个短语指代了同一主体，是该主体在文本中的两种不同的表达方式，即它们之间存在共指（Coreference）关系。就语义分析的角度而言，这两种表达方式分别蕴含了不同的语义信息。具体而言，通过“中国网友”这种表达方式，可以分析得到情绪产生的原因；而通过“他们”这种表达方式，可以得到该主体所做出的行为，即情绪引发的结果。然而，现有研究通常只将“中国网友”识别为情绪持有者，而忽略了“他们”。这就使得研究人员难以从语义层面直接识别情绪所引发的后果。在结构更为复杂的文本中，则可能对更全面、深入的情绪信息挖掘造成限制。因而，识别文本中不同表达方式之间的共指关系对于深入、全面的情绪信息分析来说非常有必要，实现这一目标的任务即为共指消解（Coreference Resolution）。此外，中文领域的情绪持有者研究缺乏相应的语料，这进一步限制了这一领域在中文环境中的研究发展。

针对上述问题，本文在中文领域的情绪持有者分析方面进行了深入研究，主要完成了以下工作：

(1) 本文以情绪原因抽取（Emotion Cause Extraction, ECE）数据集(Gui et al., 2016)和情绪驱动行为抽取（Emotion-Driven Behavior Extraction, EDBE）数据集(Sun et al., 2023)为基础，构建了一个专门针对中文新闻文本的情绪持有者抽取数据集，该数据集同时实现了情绪持有者的共指消解。

(2) 基于新构建的数据集，本文提出了一种新的模型——情绪与词性特征融合的情绪持有者抽取模型（Emotion and Part-of-speech Feature Fusion for Emotion Holder Extraction, EPF-EHE）。EPF-EHE模型结合了文本中的情绪特征、词性特征及语义特征，有效地实现了情绪持有者的抽取和共指消解。并通过对比实验和消融实验验证了模型的有效性和合理性。

2 相关工作

作为文本情绪分析的一个重要研究方向，情绪持有者的抽取在NLP领域得到了众多研究人员的关注。

早期的研究主要依赖基于规则或机器学习的方法来抽取文本中与情绪相关的语义角色。举例来说，Das and Bandyopadhyay (2010b)最初利用规则对孟加拉语博客文本中的情绪持有者进行了标注。随后，他们的工作进一步扩展，探索了使用机器学习方法来识别情绪持有者和情绪原因的可能性(Das and Bandyopadhyay, 2010a; Das and Bandyopadhyay, 2013)。Mohammad et al. (2014)则自动标注了一个与2012年美国大选相关的大型推文数据集，他们结合规则和支持向量机(Support Vector Machine, SVM)对推文进行了情绪分类，并检测了其中的情绪持有者和情绪目标。Ghazi et al. (2015)则基于FrameNet¹的情绪导向框架构建了一个数据集，该数据集标注了情绪及其持有者等语义角色，并开发了一个条件随机场(Conditional Random Fields, CRF)模型，用于检测文本中的情绪角色信息。

随着深度学习技术的突飞猛进，这一领域涌现出众多优秀的开源数据集。例如，Kim and Klinger (2018)基于古登堡计划(Project Gutenberg)²提供的文学作品，手动标注了包含情绪持有者、情绪目标等情绪角色信息的REMAN数据集，并提供了基于CRF和双向长短时记忆网络(Bidirectional Long Short-Term Memory, BiLSTM)的基线模型。Bostan et al. (2020)则创建了一个包含5000个新闻标题的语料库GoodNewsEveryone，这些标题经过众包标注，涵盖了情绪、情绪持有者、情绪线索、情绪目标、情绪原因以及读者对标题情绪的感知，同时提供了一个基于CRF和BiLSTM的基线模型。在上述研究的基础上，Campagnano et al. (2022)提出了一个专注于情绪语义角色标注的评估框架SRL4E，通过采用统一的标注方案来整合多个已标注情绪和语义角色的数据集。

在以BERT(Devlin et al., 2019)为代表的预训练语言模型在NLP各项任务中取得显著成功后，研究人员开始利用这些模型来抽取情绪角色信息。例如，Wegge et al. (2022; 2023)基于BERT等预训练语言模型微调构建了多任务学习模型，以同时抽取各类情绪角色信息。

3 情绪持有者抽取数据集

正如第2节中所述，目前已经有国外的研究人员推出了解决情绪持有者抽取问题的数据集。然而，这些数据集并未考虑文本中存在的情绪持有者表达的共指问题。此外，情绪持有者的抽取在中文领域尚缺乏相应数据集的支撑。为了解决这些问题，本文在ECE数据集和EDBE数据集的基础上，构建了针对中文文本的情绪持有者抽取数据集，该数据集同时解决了文本中存在的情绪持有者的共指消解问题。

本节，将详细介绍本文工作所构建的情绪持有者抽取数据集，主要包括数据源的选取与数据预处理、数据标注与数据集的统计。

3.1 数据源的选取与数据预处理

本文基于ECE数据集展开研究，保留了数据集中的情绪类别和情绪线索，并对每个情绪线索所对应的情绪持有者进行详细标注。ECE数据集采用了Ekman (1992)提出的情绪分类标准，该标准将基本情绪划分为愤怒(Anger)、厌恶(Disgust)、恐惧(Fear)、喜悦(Happiness)、悲伤(Sadness)和惊讶(Surprise)六类，这一分类方案在W3C情绪标记语言中被称为“Big6”方案³。ECE数据集选取2013年至2015年间发布的新浪新闻⁴文本数据进行标注，共获得2105条标注数据。

考虑到ECE数据集的数据量较小，本文进一步选取了THUCNews数据集⁵中“社会”类别下的新闻文本进行扩展。首先，对数据进行清洗，移除如“\u3000”、“ ”等特殊字符。随后，通过识别“。”、“？”、“！”等可视为句子终结的标点符号来对文本进行分句。在得到分句后的文本基础上，本文使用基于BERT的情绪分类模型，该模型是基于ECE数据集训练的，用以对分句后的文本进行情绪分类。若某个句子被判断属于特定情绪类别，便将其选入待标注数据集。

¹<http://framenet.icsi.berkeley.edu/>

²<https://www.gutenberg.org/>

³<https://www.w3.org/TR/emotion-voc/xml#big6>

⁴<http://news.sina.com.cn/society/>

⁵<http://thuctc.thunlp.org/>

3.2 数据集的标注

在数据标注阶段，本文使用开源的数据标注工具doccano⁶进行数据标注。

具体而言，本文首先增加了一个用于标注情绪持有者的标签“Holder”。为了保证标注的情绪持有者尽可能包含更加全面的信息，本文所标注的情绪持有者应包含其所有的定语，如“张三的好朋友李四”应标注这一完整表达，而不能只标注“李四”。同时，增加了“情绪-持有者 (Emotion-Holder)”和“共指 (Coreference)”两个关系标签来解决一段文本存在多个情绪线索的情况下而导致的情绪线索与情绪持有者表达无法准确对应的问题。其中，“Emotion-Holder”为连接情绪线索及与其在语法上存在直接关系的持有者表达的关系，对于每一个情绪线索，虽然它的情绪持有者在文中可能存在多种表达形式，但是本文只将与其在语法上存在直接关系的情绪持有者表达进行标注。“Coreference”则表示共指关系，用于连接情绪持有者的不同表达方式。

为了更加清晰地展示本文的标注方案，本文给出了一个使用doccano对数据进行情绪持有者标注的示例。在该示例中，“感动”一词作为情绪线索代表了一种“喜悦”的情绪，因而其标签为“Happiness”。这一情绪的情绪持有者在文中出现了两处表达方式，分别是“中国网友”和“他们”，因而将这两种表达标注为“Holder”。在关系上，文中出现的第一个“中国网友”是“感动”的宾语，他们之间在语法上存在直接关系，因而将文中的“感动”与“中国网友”之间标注为“Emotion-Holder”关系。文中情绪持有者的两个表达之间则存在共指关系，为了统一数据标注的标准，本文将该情绪持有者的第一个表达之后的其余表达均与第一个表达之间标注为共指关系，即“他们”与“中国网友”之间的关系为“Coreference”。

在进行标注后，本文以情绪线索为基础来形成数据。具体而言，如果同一段文本中有多个情绪线索，则将各情绪线索与其对应的情绪持有者表达单独分离出来构成一条数据。

3.3 数据集统计信息

为了理解情绪持有者抽取任务的数据特性，本文对所标注的数据集进行了详尽的统计与分析。这些统计结果不仅揭示了数据集中的基础数量特征，还探究了情绪线索与持有者之间的关系，以及不同情绪类别下的持有者特征分布。

首先，表1展示了情绪持有者抽取数据集的基础统计信息。数据集共包含3971条数据，其中每条数据可能包含一个或多个情绪持有者的表达。从表中可以看出，大多数数据（2008条）包含一个情绪持有者表达，而包含两个或更多持有者的数据则相对较少。此外，平均每条数据包含1.8个情绪持有者的共指表达。

项目	统计值
数据量	3971
有1个情绪持有者表达的数据量	2008
有2个情绪持有者表达的数据量	1135
有3个情绪持有者表达的数据量	509
有4个及以上情绪持有者表达的数据量	319
每条数据的平均持有者表达数	1.8

Table 1: 情绪持有者抽取数据集的统计信息

进一步地，本文对数据集中不同情绪类别的分布情况进行了统计。如表2所示，恐惧和悲伤是数据集中出现频率最高的两种情绪类别，分别占据了总情绪数的25.66%和24.98%。这表明在所研究的文本数据中，这两种情绪是相对较为突出的。相比之下，愤怒、厌恶和喜悦等情绪类别则占据了中等比例，而惊讶情绪的出现频率最低，仅占2.74%。这表明了本文构建的数据集通其他主流数据集一样(Demszky et al., 2020)，也存在数据不平衡的问题。进一步来讲，文章构建的数据集中负面情绪较多，这与其数据来源有密切关系——负面情绪往往更容易引发一些过激行为，从而构成新闻事件。

⁶<https://github.com/doccano/doccano>

情绪类别	数量	百分比
恐惧	1019	25.66%
悲伤	992	24.98%
愤怒	714	17.98%
厌恶	487	12.27%
喜悦	650	16.37%
惊讶	109	2.74%

Table 2: 情绪持有者抽取数据集中的情绪类别统计

4 方法

针对中文情绪持有者抽取问题，本文提出EPF-EHE模型。该模型将情绪持有者的抽取任务看做一个序列标注任务。具体而言，针对文本中的每一个情绪线索，抽取其情绪持有者及情绪持有者的所有共指表达形式。在模型设计上，首先通过语义特征编码层、情绪特征编码层和词性特征编码层来分别获取文本中每个字符的语义特征、情绪特征以及词性特征；然后通过特征融合层将得到的上述三个特征进行融合；接着利用BiLSTM将特征融合层得到的融合特征进行进一步的编码；最后利用CRF对特征进行解码，得到序列标注的结果。

EPF-EHE模型的总体架构如图2所示。接下来，本节将对该模型进行详细介绍。

4.1 语义编码层

在自然语言处理任务中，一项基础性的工作就是将文本转化成向量的表示形式，以便于后续对文本数据进行分析与计算。具体而言，对于长度为 n 的输入 $X = [x_1, x_2, \dots, x_n]$ ，该模块就是将组成 X 的每个 x_i 转变成一个带有上下文语义信息的向量 v_i ，即将 X 转变为 $V = [v_1, v_2, \dots, v_n]$ ，其中 $i \in [0, n]$ 。

众所周知，利用最原始的独热编码进行编码后得到的向量不包含语义信息，而利用像Word2Vector这样的静态词向量进行编码，得到的向量虽然包含一定的语义信息，但是这些向量中无法携带上下文信息。因而，本文采用基于Transformer的预训练语言模型作为文本的编码器。为了便于与其他模型进行比较，本文采用经典的基础版BERT_{base}模型作为编码器。

BERT是一种双向编码表示的预训练语言模型，它在大量文本数据的基础上，基于掩码语言建模和下一句预测这两个无监督的预训练任务来学习语言表示。BERT对文本进行编码，需要在文本的开头加一个句子分类标记([CLS])，在文本的结尾加一个句子分隔标记([SEP])。因而，使用BERT对 X 进行编码，需要首先在 X 的开头加一个句子分类标记 $x_{[CLS]}$ ，在 X 的结尾加一个句子分隔标记 $x_{[SEP]}$ ，即 $X = [x_{[CLS]}, x_1, x_2, \dots, x_n, x_{[SEP]}]$ 。相应地，编码得到的 $V = [v_{[CLS]}, v_1, v_2, \dots, v_n, v_{[SEP]}]$ 。对于BERT_{base}，它输出的向量维度为768，即 V 中每个 v_i 的维度为 1×768 ，则 V 的维度为 $(n + 2) \times 768$ 。

4.2 情绪特征编码层

一段文本中可能存在多个情绪线索，要针对指定的长度为 m 的情绪线索 $C_e = [x_{e1}, x_{e2}, \dots, x_{em}]$ 来抽取情绪的持有者，那么就需要以某种方式将情绪线索 C_e 的信息传递到本文的情绪持有者抽取模型中。

一种比较简单的方法就是直接将 C_e 与 X 拼接得到一个新的输入 X' ，则 X' 的一种可能取值为 $[x_{[CLS]}, x_{e1}, x_{e2}, \dots, x_{em}, x_{[SEP]}, x_1, x_2, \dots, x_n, x_{[SEP]}]$ ，对于图1中的例1而言 X' 对应的文本即为“[CLS]感动[SEP]伊拉克选手的境况感动了中国网友，除了发起向伊拉克队员的捐助，他们还呼吁为伊拉克队颁发一枚独特的‘金牌’。[SEP]”。然而，这就存在这样一个问题：一段文本中可能会有多个相同的情绪线索，而这些情绪所对应的持有者是不同的。例如，对于这样的一段文本：“警察制服了情绪激动的绑匪，围观群众激动地鼓起掌来”，其中“激动”出现了两次，然而这两个“激动”所表达的情绪其实是不同的。第一个“激动”表达的是一种“愤怒”的情绪，这一情绪的持有者是“情绪激动的绑匪”。而第二个“激动”则表达了一种“快乐”的情绪，这一情绪的持有者则是“围观群众”。显然，这种情况下，如果直接采用拼接的方法构成新的输入 X' ，则容易对模型造成困惑，最终无法得到正确的结果。

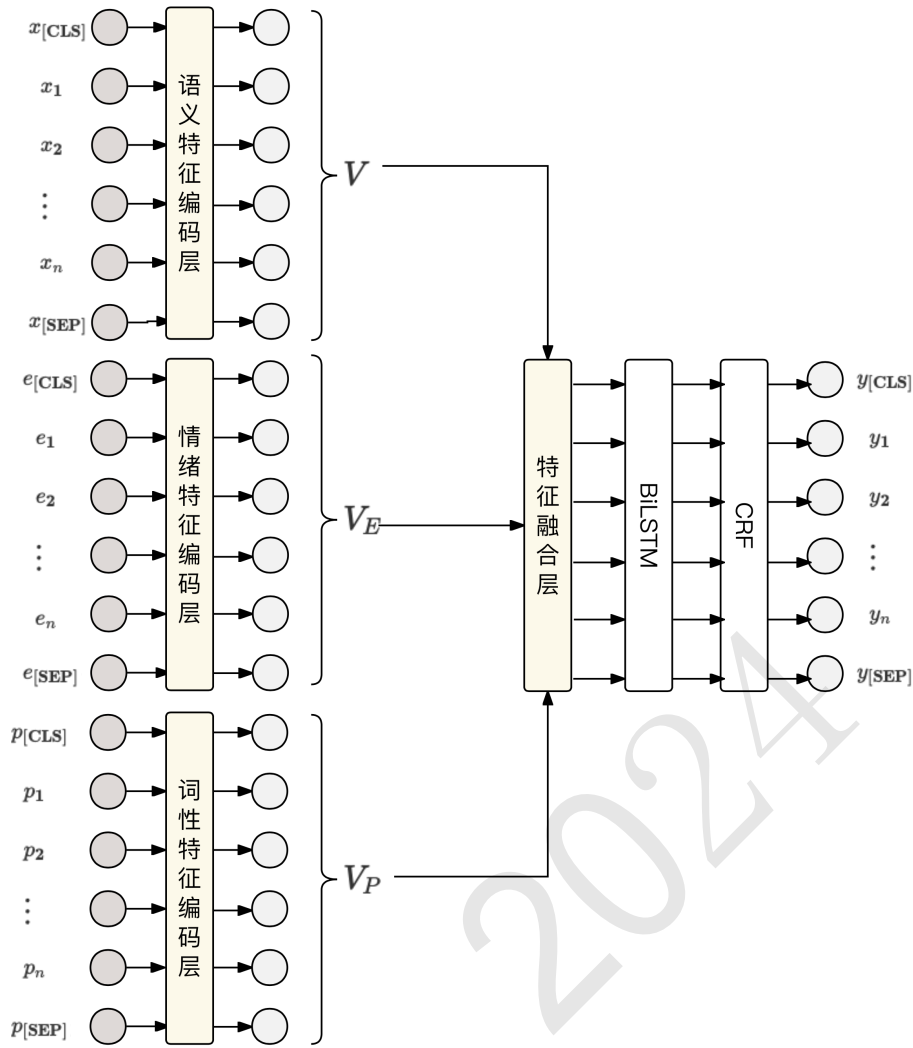


Figure 2: EPF-EHE模型框架图

为了解决这一问题，本文在情绪持有者抽取模型中增加了对情绪线索位置的编码层，即情绪特征编码层。具体而言，就是将指定的情绪线索在原始文本中对应的位置标为1，其他位置则标为0，进而得到情绪编码序列 $E = \{e_{[SEP]}, e_1, e_2, \dots, e_n, e_{[SEP]}\}$ ，其中 $e_i \in \{0, 1\}$ 。由于[CLS]与[SEP]肯定不会属于任意一个情绪线索的组成部分，因而 $e_{[SEP]}$ 与 $e_{[CLS]}$ 默认为0。对于情绪线索 $C_e = \{x_{e1}, x_{e2}, \dots, x_{em}\}$ ， E 中的元素 e_i 计算方法如式 (1) 所示。

$$e_i = \begin{cases} 1, & i \in [e1, em] \\ 0, & \text{其他} \end{cases} \quad (1)$$

为了将情绪特征与文本 X 经由语义编码层编码得到的语义特征 V 相融合，本文通过增加一个嵌入层 $Embedding_E$ 将 E 编码成与 V 形状相同的向量 V_E ，如式 (2) 所示。

$$V_E = Embedding_E(E) \quad (2)$$

嵌入层本质上是一种词典，其中保存了它所要编码的状态所对应的向量。由于 p_{ei} 只有0和1两种状态，而且经过嵌入层编码 $Embedding_E$ 得到的情绪特征向量 V_E 需要与语义特征向量 V 一致，因而，本文对情绪特征进行编码的嵌入层 $Embedding_E$ 的形状为 2×768 。根据 V_E 的定义可知，其长度为 $n + 2$ ，而 $Embedding_E$ 的形状为 2×768 ，则 V_E 的形状为 $(n + 2) \times 768$ 。

嵌入层的参数是可训练的，因而，在训练前需要先对 $Embedding_E$ 进行初始化。在训练过程中，模型可以动态更新 $Embedding_E$ 的参数，最终达到对情绪特征进行编码的效果。

4.3 词性特征编码层

情绪持有者，从定义上来讲就是名词性的词语或短语。从本文的研究目标来看，本文抽取得到的情绪持有者可能是包含了完整定语的人物、机构，也可能是代词。而情绪持有者在文本的语言表达上定语则可能是形容词等其他词性的词。如果能够将文本中的命名实体和词性信息作为特征融合到情绪持有者抽取模型中，很可能将有助于模型效果的提升。

这涉及自然语言处理领域的命名实体识别和词性标注这两个基础性任务。目前，已经有很多比较成熟的工具可以帮助用户来处理这方面的工作，像LAC(Jiao et al., 2018)⁷、LTP(Che et al., 2021)⁸、HarvestText⁹以及HanLP(He and Choi, 2021)¹⁰等，都是在中文自然语言处理领域非常优秀的开源工具。上述提到的开源工具，多数都是将命名实体识别与词性标注看做两个独立的任务，如果采用这样的工具，则需要对文本中的命名实体和词性进行分别处理。然而，从命名实体的定义上来讲，它就是名词性的短语，将它与词性标注分开来处理在本文任务中显然是没有必要的。百度公司开发的LAC则不存在这样的问题。LAC将命名实体识别与词性标注整合为一个任务，其将命名实体类别也视为词性，采用统一的接口来实现。LAC的包含命名实体类别在内的词性类别标签如表3所示。

序号	词性标签	含义	序号	词性标签	含义	序号	词性标签	含义
0	n	普通名词	9	ORG	机构名	18	an	名形词
1	nz	其他专名	10	nw	作品名	19	r	代词
2	a	形容词	11	vn	名动词	20	d	副词
3	m	数量词	12	ad	副形词	21	p	介词
4	c	连词	13	q	量词	22	w	标点符号
5	PER	人名	14	u	助词	23	TIME	时间
6	f	方位名词	15	LOC	地名	24	o	其他符号
7	v	普通动词	16	s	处所名词			
8	xc	其他虚词	17	vd	动副词			

Table 3: LAC的词性类别标签信息表

本文利用LAC对文本进行词性标注，得到的是该文本分词后的词语（或短语）列表以及相应的词性标签列表。要想完成对词性信息的编码，需要先对每个词性类别进行编号。本文对每一个词性类别都进行了编号，如表3的“序号”列所示。考虑到EPF-EHE模型在语义编码层对数据进行处理时，需要分别在文本的开头和结尾添加[CLS]和[SEP]这两个符号，同时，对文本进行补齐（Padding），这三种符号不属于原始文本的组成部分，也就不具备任何的词性。为了处理这种情况，本文在表3中增加了标签“o”，代表上述提到的这三种情况。

为了直观地展示使用LAC对文本进行词性标注及相应的编号结果，本文同样以图1中的示例为例。在对中文文本中的词汇进行词性标注时，通常需要先对文本进行分词，然后再以词为单位进行词性标注。然而，BERT对中文文本进行处理并不是基于词汇这一级别的，而是基于字符级别。也就是说，BERT无法直接处理“伊拉克”这样的词汇，而需要先将“伊拉克”分隔为“伊”、“拉”和“克”三个字，然后分别处理这三个字。在示例中，“伊拉克”的词性标签为“LOC”，则本文将它分隔为单个字符后，使组成它的“伊”、“拉”、“克”的词性标签与它本身保持一致。对示例进行词性标注得到的其他词汇，本文均采用相同的处理方法，得到的结果如表4所示。

表4中“编号”列可组成文本X的词性序列 $P = [p_{[CLS]}, p_1, p_2, \dots, p_n, p_{[SEP]}]$ 。与情绪特征编码层相同，想要完成对词性特征的编码，同样需要一个嵌入层 $Embedding_P$ 。表3中一共列举了25中词性类型，那么 $Embedding_P$ 的形状相应地应该为 25×768 。

通过使用嵌入层 $Embedding_P$ 对词性词性序列POS进行编码，可以得到X的词性特征编码结果 V_P ，如式（3）所示。POS的长度显然为 $n + 2$ ，而 $Embedding_P$ 的形状为 25×768 ，则 V_P 的形状

⁷<https://github.com/baidu/lac>

⁸<https://github.com/HIT-SCIR/ltp>

⁹<https://github.com/blmoistawinde/HarvestText>

¹⁰<https://github.com/hankcs/HanLP>

词汇	词性标签	词性编号	词汇	词性标签	词性编号	词汇	词性标签	词性编号
[CLS]	o	24	了	p	21	为	v	7
伊	LOC	15	发	v	7	伊	ORG	9
拉	LOC	15	起	v	7	拉	ORG	9
克	LOC	15	向	p	21	克	ORG	9
选	n	0	伊	LOC	15	队	ORG	9
手	n	0	拉	LOC	15	颁	v	7
的	u	14	克	LOC	15	发	v	7
境	n	0	队	n	0	一	m	3
况	n	0	员	n	0	枚	m	3
感	v	7	的	u	14	独	a	2
动	v	7	的	vn	11	特	a	2
了	u	14	助	vn	11	的	u	14
中	LOC	15	,	w	22	‘	w	22
国	LOC	15	他	r	19	金	n	0
网	n	0	们	r	19	牌	n	0
友	n	0	还	d	20	’	w	22
,	w	22	呼	v	7	。	w	22
除	p	21	吁	v	7	[SEP]	o	24

Table 4: 文本的字符级词性标注示例

为 $(n + 2) \times 768$ 。

$$V_P = \text{Embedding}_P(P) \tag{3}$$

与 Embedding_E 相同，本文在训练前需要先对 Embedding_P 进行初始化。在训练过程中，模型同样的可以动态更新 Embedding_P 的参数，最终达到对词性特征进行编码的效果。

4.4 特征融合层

对长度为 n 的文本 X ，分别经过语义编码层、情绪特征编码层和词性特征编码层的编码之后，得到 V 、 V_E 和 V_P 这三个形状均为 $(n + 2) \times 768$ 的特征向量。参考BERT在嵌入层对符号嵌入、位置嵌入和段落嵌入这三种嵌入的处理方法，将得到的三个特征向量逐元素相加，从而得到融合后的特征向量 V_c ，如式 (4) 所示。

$$V_c = V + V_E + V_P \tag{4}$$

4.5 基于融合特征向量的情绪持有者抽取

在经过语义编码层、情绪特征编码层和词性特征编码层的编码，并通过特征融合层对三种编码后的特征向量进行融合后，得到了融合了上下文语义信息、情绪信息、词性信息的特征向量 V_c 。接下来，本文的主要任务就是利用这样的特征向量得到最终的抽取结果。

本文将情绪持有者的抽取任务看做一个序列标注的任务来完成。即对于输入 $X = [x_{[CLS]}, x_1, x_2, \dots, x_n, x_{[SEP]}]$ ，需要得到序列标签 $Y = [y_{[CLS]}, y_1, y_2, \dots, y_n, y_{[SEP]}]$ 。这就涉及对组成 X 的字符 x_i 进行标注的问题，本文采用经典的“BIO”标注方案来完成标注。具体而言，就是将 X 中组成情绪持有者的字符标注成“B”或者“I”，将其他字符标注为“O”。对于组成情绪持有者的字符，如果它是相应的情绪持有者的第一个字符，则将其标注为“B”，否则标注为“I”。

为了更加清晰地阐述本文的标注方案，本文给出了针对图1所示的示例中根据“感动”这一情绪线索所抽取到的情绪持有者的标注结果，如表5所示。

在得到融合特征向量 V_c 以后，是难以直接得到上述这种标注结果的。为了进一步捕获 V_c 的上下文语义信息，本文在特征融合层之后增加了一个BiLSTM层来对特征向量进行进一步的编码。BiLSTM能够从两个方向捕获序列的上下文依赖关系，能够进一步增强每一个字符所对应特征向量的语义表示能力。在经过BiLSTM进行进一步的编码后，得到 V_c 的隐藏向量表示 H ，如式 (5) 所示。

$$H = \text{BiLSTM}(V_c) \tag{5}$$

字符	标签	字符	标签	字符	标签	字符	标签	字符	标签	字符	标签
[CLS]	O	感	O	了	O	的	O	为	O	独	O
伊	O	动	O	发	O	捐	O	伊	O	特	O
拉	O	了	O	起	O	助	O	拉	O	的	O
克	O	中	B	向	O	,	O	克	O	'	O
选	O	国	I	伊	O	他	B	队	O	金	O
手	O	网	I	拉	O	们	I	颁	O	牌	O
的	O	友	I	克	O	还	O	发	O	,	O
境	O	,	O	队	O	呼	O	一	O	。	O
况	O	除	O	员	O	吁	O	枚	O	[SEP]	O

Table 5: 情绪持有者抽取的标注示例

接下来, 需要对经过上述的一系列编码得到向量表示进行解码。本节中, 采用CRF来对 H 进行解码。不同于简单地为每个词独立地选择标签, CRF考虑了标签序列的约束, 优化整个序列的标签选择。这意味着模型在为一个词选择标签时, 会考虑到整个句子中所有词的标签, 确保最终的标签序列在全局上是最优的。经过运用CRF进行解码, 最终得到了对输入序列 X 的预测, 如式 (6) 所示。

$$\hat{Y} = \text{CRF}(H) \quad (6)$$

式中, \hat{Y} 则表示模型对输入序列 X 的预测, 其中 $\hat{Y} = \{\hat{y}_{[\text{CLS}]}, \hat{y}_1, \hat{y}_2, \dots, \hat{y}_n, \hat{y}_{[\text{SEP}]}\}$ 。

通常情况下, 在模型训练之初, 预测值 \hat{Y} 与标签 Y 之间会存在较大的差异。而本文训练模型的目的就是减少 \hat{Y} 与 Y 之间的差异。

5 实验

本节, 将介绍本章工作所构建的数据集, 同时将介绍文章工作的实验设置、实验结果及分析等内容。

5.1 实验设置

本章在语义编码层采用BERT针对汉语的基础版本, 即“bert-base-chinese”。所有实验均使用Tesla V100-SXM2-16GB GPU进行加速。模型采用PaddlePaddle框架来实现。在训练过程中, 采用交叉熵损失作为损失函数, 使用AdanW作为优化器来最小化损失。在训练过程中, 取验证集上 F_1 值最高的模型参数作为最佳模型参数。

在超参数方面, BERT的学习率设置为 1×10^{-5} , 其他部分的学习率则设置为 1×10^{-3} 。实验中未设置最大训练步数, 而是采用早停策略来控制训练过程。具体而言, 每训练20步即使用验证集进行一次验证, 若模型在连续20次验证中未能取得更好的效果, 则终止训练。

5.2 评价指标

本章采用精确率 (Precision, P)、召回率 (Recall, R) 和 F_1 值作为实验的评价指标。对于模型预测的一种情绪持有者表达, 如果能够与人工标注的情绪持有者表达完全匹配上, 则为真阳性 (True Positive, TP); 如果模型预测的情绪持有者表达无法与人工标注的情绪持有者表达匹配上, 则为假阳性 (False Positive, FP); 如果人工标注的情绪持有者表达没有被模型预测到, 则为假阴性 (False Negative, FN)。则精确率、召回率和 F_1 值的计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (7)$$

$$R = \frac{TP}{TP + FN} \quad (8)$$

$$F_1 = \frac{2 \cdot P \cdot R}{P + R} \quad (9)$$

5.3 基线模型

首先，本章选取如下模型作为基线模型：

BERT：该模型是以BERT为基础进行文本序列标注任务的最基础的模型。该模型以BERT作为编码器，以Softmax作为解码器。是HuggingFace¹¹与PaddlePaddle¹²官方提供的基线模型。

BERT-CRF：该模型以BERT编码器，以CRF作为解码器，被广泛应用于文本序列标注任务中(Souza et al., 2019)。

BERT-BiLSTM-CRF：该模型是各类文本序列标注任务中的经典模型。该模型以BERT和BiLSTM作为编码器，以CRF作为解码器，在文本的各类序列标注任务中均取得了较好的效果，并被众多研究作为基线模型(Bostan et al., 2020)。

SRL4E：该模型是Campagnano et al. (2022)在其情绪角色标注统一评估框架中提出的模型。对于文本 X 和情绪线索 E ，SRL4E构造了新的文本序列 $X' = [\text{CLS}]X[\text{SEP}]E[\text{SEP}]$ ，将 X' 利用BERT进行编码后，再进一步利用两层的BiLSTM进行语义编码，最终利用Softmax进行解码。

Code4Struct：该方法是Wang et al. (2023)针对事件论元抽取任务所提出的基于大预言模型(Large Language Model, LLM)的方法。该方法利用LLM强大的文本理解与推理能力，将自然语言处理任务转化为结构化的代码，进而完成将文本转化为结构化数据的任务。

ChatGPT 3.5¹³：该模型是OpenAI公司推出的生成式大语言模型。本文通过对话形式，直接就情绪持有者抽取任务对ChatGPT进行询问，从而得到答案，并通过文本匹配的方式进行各项指标的统计与计算。

5.4 结果与分析

(1) 情绪持有者抽取的实验结果与分析

表6详细展示了在情绪持有者抽取任务中，EPE-EHE与其他基线模型比较的结果。从表中可以看出，EPE-EHE在精确率、召回率和 F_1 分数这三个关键指标上均取得了最佳表现。

具体来说，EPE-EHE的精确率达到了82.39%，这意味着在模型预测为情绪持有者的样本中，有82.39%是真正的情绪持有者。这一结果明显高于其他基线模型。其中，BERT作为表现次优的模型，其精确率为81.05%，相较于EPE-EHE低了1.34个百分点。

在召回率方面，EPE-EHE同样表现出色，达到了87.66%。这意味着在实际为情绪持有者的样本中，有87.66%被EPE-EHE正确识别。在基线模型中，经典的BERT-BiLSTM-CRF模型的召回率反而更接近EPE-EHE，这主要是因为该模型未考虑情绪线索，这也导致其在精确率上表现最差。

需要注意的是，基于LLM的Code4Struct和ChatGPT 3.5在各项指标上的效果均不理想。虽然主流的LLM模型在文本的语义理解及推理上取得了令人惊叹的效果，然而，在大规模的数据处理上，特别是信息抽取的相关任务中，还存在一定的缺陷。在本研究中，通过试验发现，在进行多轮对话的过程中，LLM比较严重地依赖提示信息，回答的结果不够稳定，在一些问题上会自行扩展一些额外的信息，容易出现答非所问的情况，这些都导致了其在任务中表现不佳的情况。

综合精确率和召回率，EPE-EHE的 F_1 分数达到了84.94%，这一综合指标同样高于其他所有基线模型。在基线模型中，BERT和BERT-BiLSTM-CRF分别在精确率和召回率上达到了接近EPE-EHE的水平，然而他们分别在召回率和精确率上取得了最差的效果，因而综合表现与EPE-EHE之间存在差距。而SRL4E的各项指标表现得中规中矩，Code4Struct与ChatGPT 3.5则在各项指标中均效果不佳。因而EPE-EHE在综合性能上在基线模型中取得了相对领先的效果。

(2) 消融实验

消融实验的目的是评估模型各个组成部分的贡献。通过依次移除模型的某一部分（例如，特定的特征或层），可以观察到模型性能的变化，从而理解该部分对模型整体性能的影响。在

¹¹<https://github.com/huggingface/transformers>

¹²<https://github.com/PaddlePaddle/PaddleNLP>

¹³<https://chatgpt.com>

模型	精确率	召回率	F_1
BERT	81.05%	78.67%	79.84%
BERT-CRF	74.42%	81.84%	77.95%
BERT-BiLSTM-CRF	66.12%	86.52%	74.96%
SRL4E	74.56%	81.84%	78.03%
Code4Struct	22.26%	24.70%	23.42%
ChatGPT 3.5	13.98%	19.63%	16.33%
EPF-EHE	82.39%	87.66%	84.94%

Table 6: 情绪持有者抽取的实验结果

情绪直接持有者抽取任务和情绪持有者抽取任务中，EPF-EHE模型的消融实验结果展示了移除不同组件后模型性能的变化。

V_E	V_p	BiLSTM	CRF	精确率	召回率	F_1
+	+	+	+	82.39%	87.66%	84.94%
-	+	+	+	78.86%	83.79%	81.25%
+	-	+	+	81.03%	82.62%	81.82%
+	+	-	+	70.80%	78.13%	74.28%
+	+	+	-	70.85%	85.94%	77.67%

Table 7: 情绪持有者抽取消融实验结果

表7展示了消融实验结果。消融实验结果表明，情绪特征编码、词性特征编码、BiLSTM层和CRF层都对EPF-EHE模型在情绪持有者抽取任务中的性能起到了积极作用。特别是BiLSTM层的移除导致了最为显著的性能下降，这说明上下文信息对于情绪持有者的抽取至关重要。情绪特征编码和词性特征编码的贡献也不容忽视，它们通过提供关于情绪线索位置和词性的信息，帮助模型更精准地定位和理解情绪持有者。尽管CRF层的移除对性能的影响相对较小，但它在优化整个序列的标签选择中仍然发挥着重要作用，特别是在保证标签序列全局一致性方面。

这些结果证明了模型不同组件在处理复杂的情绪持有者抽取任务时的综合作用。情绪特征编码和词性特征编码通过提供关键的文本信息，增强了模型对情绪持有者的识别能力；BiLSTM层通过捕获句子的上下文依赖关系，增强了模型对情绪持有者与情绪线索之间关系的理解；CRF层则在序列标注任务中优化了标签序列的全局一致性。

6 总结

本文探讨了文本情绪分析中一个重要但经常被忽视的组成部分——情绪持有者的识别和抽取。情绪持有者是指在文本中显示出特定情绪的实体。准确识别这些实体对于深入理解文本中的情绪信息至关重要。

首先，本文分析了现有研究中的主要挑战，特别是情绪持有者共指现象的处理问题以及针对中文文本的资源匮乏问题。为了应对这些挑战，本文构建了一个专门针对中文文本的情绪持有者抽取数据集。该数据集不仅丰富了中文情绪分析的资源，还通过精确的标注解决了共指现象的问题。

其次，本文提出了一种新的情绪持有者抽取模型，该模型综合了语义特征、情绪特征和词性特征，通过这种融合方法显著提高了情绪持有者抽取的准确性。模型采用BERT作为基础编码器，利用BiLSTM来进一步捕获情绪表达的上下文关系，最后通过CRF层优化标注序列，提高了整体的标注效果。

在实验部分，通过与多个基线模型的对比，验证了所提模型在多个性能指标上的优越性。此外，本文通过消融实验，进一步证明了模型中各个组件的有效性和必要性。

下一步，将以情绪持有者抽取为基础，对文本的情绪分析进行更深层次的研究。

参考文献

- Laura Bostan, Evgeny Kim, and Roman Klinger. 2020. GoodNewsEveryone: A Corpus of News Headlines Annotated with Emotions, Semantic Roles, and Reader Perception. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 1554–1566.
- Cesare Campagnano, Simone Conia, and Roberto Navigli. 2022. SRL4E – Semantic Role Labeling for Emotions: A Unified Evaluation Framework. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 4586–4601, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Wanxiang Che, Yunlong Feng, Libo Qin, and Ting Liu. 2021. N-LTP: An Open-source Neural Language Technology Platform for Chinese. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 42–49, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Dipankar Das and Sivaji Bandyopadhyay. 2010a. Finding Emotion Holder from Bengali Blog Texts-An Unsupervised Syntactic Approach. In *Proceedings of the 24th Pacific Asia Conference on Language, Information and Computation*, pages 621–628.
- Dipankar Das and Sivaji Bandyopadhyay. 2010b. Labeling Emotion in Bengali Blog Corpus – A Fine Grained Tagging at Sentence Level. In Sarmad Hussain, Virach Sornlertlamvanich, and Hammam Riza, editors, *Proceedings of the Eighth Workshop on Asian Language Resources*, pages 621–628, Beijing. Coling 2010 Organizing Committee.
- Dipankar Das and Sivaji Bandyopadhyay. 2013. Emotion Co-referencing-Emotional Expression, Holder, and Topic. *International Journal of Computational Linguistics & Chinese Language Processing*, 18(1):79–98.
- Dorottya Demszky, Dana Movshovitz-Attias, Jeongwoo Ko, Alan Cowen, Gaurav Nemade, and Sujith Ravi. 2020. GoEmotions: A Dataset of Fine-Grained Emotions. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4040–4054, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.
- Keyang Ding, Jing Li, and Yuji Zhang. 2020. Hashtags, Emotions, and Comments: A Large-Scale Dataset to Understand Fine-Grained Social Emotions to Online Topics. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1376–1382, Stroudsburg, PA, USA. Association for Computational Linguistics.
- R. J. Dolan. 2002. Emotion, Cognition, and Behavior. *Science*, 298(5596):1191–1194, 11.
- Paul Ekman. 1992. An argument for basic emotions. *Cognition and Emotion*, 6(3-4):169–200, 5.
- Diman Ghazi, Diana Inkpen, and Stan Szpakowicz. 2015. Detecting Emotion Stimuli in Emotion-Bearing Sentences. In *Computational Linguistics and Intelligent Text Processing*, volume 9042, pages 152–165. Springer International Publishing, Cham.
- Lin Gui, Dongyin Wu, Ruifeng Xu, Qin Lu, and Yu Zhou. 2016. Event-Driven Emotion Cause Extraction with Corpus Construction. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 1639–1649, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Han He and Jinho D. Choi. 2021. The Stem Cell Hypothesis: Dilemma behind Multi-Task Learning with Transformer Encoders. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 5555–5577, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Zhenyu Jiao, Shuqi Sun, and Ke Sun. 2018. Chinese Lexical Analysis with Deep Bi-GRU-CRF Network. *arXiv preprint arXiv:1807.01882*, 7.

- Evgeny Kim and Roman Klinger. 2018. Who Feels What and Why? Annotation of a Literature Corpus with Semantic Roles of Emotions Title and Abstract in German. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 1345–1359, Santa Fe, New Mexico, USA.
- Jasy Liew, Suet Yan, Howard R Turtle, and Elizabeth D Liddy. 2016. EmoTweet-28: A Fine-Grained Emotion Corpus for Sentiment Analysis. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 1149–1156.
- Saif Mohammad, Xiaodan Zhu, and Joel Martin. 2014. Semantic Role Labeling of Emotions in Tweets. In *Proceedings of the 5th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, pages 32–41, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Fábio Souza, Rodrigo Nogueira, and Roberto Lotufo. 2019. Portuguese Named Entity Recognition using BERT-CRF. *arXiv preprint arXiv:1909.10649*, 9.
- Yawei Sun, Saike He, Xu Han, and Ruihua Zhang. 2023. A New Model for Emotion-Driven Behavior Extraction from Text. *Applied Sciences*, 13(15):8700, 7.
- Xingyao Wang, Sha Li, and Heng Ji. 2023. Code4Struct: Code Generation for Few-Shot Event Structure Prediction. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 3640–3663, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Maximilian Wegge and Roman Klinger. 2023. Automatic Emotion Experiencer Recognition. In *Proceedings of the 3rd Workshop on Computational Linguistics for the Political and Social Sciences*, pages 1–7.
- Maximilian Wegge, Enrica Troiano, Laura Ana Maria Oberlaender, and Roman Klinger. 2022. Experiencer-Specific Emotion and Appraisal Prediction. In *Proceedings of the Fifth Workshop on Natural Language Processing and Computational Social Science (NLP+CSS)*, pages 25–32, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Duzhen Zhang, Feilong Chen, and Xiuyi Chen. 2023. DualGATs: Dual Graph Attention Networks for Emotion Recognition in Conversations. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 7395–7408, Stroudsburg, PA, USA. Association for Computational Linguistics.